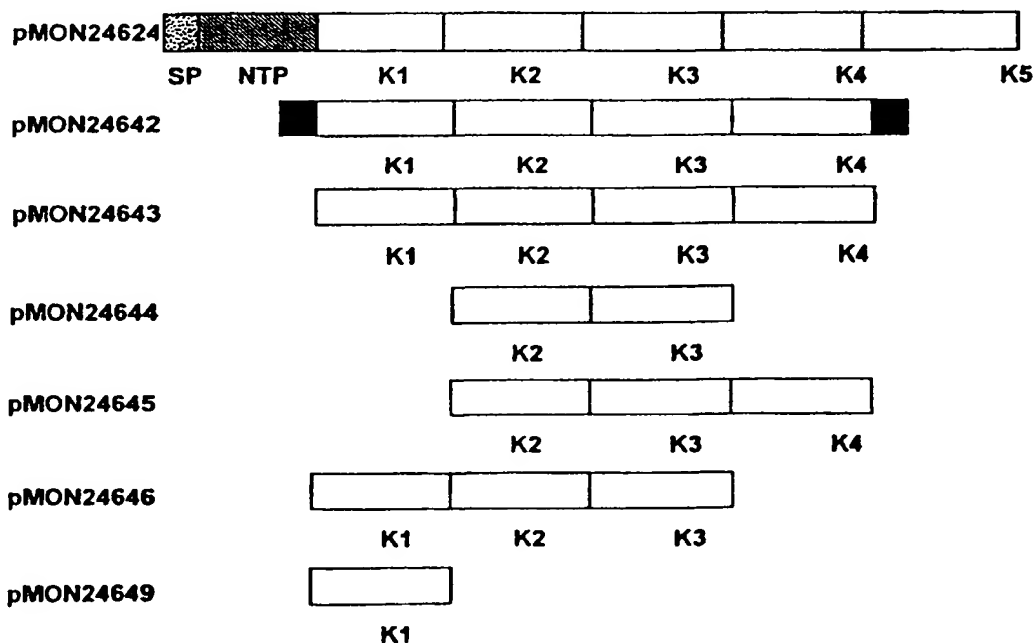




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C12N 9/68, C07K 1/113	A2	(11) International Publication Number: WO 99/35248 (43) International Publication Date: 15 July 1999 (15.07.99)
(21) International Application Number: PCT/US99/00048 (22) International Filing Date: 11 January 1999 (11.01.99) (30) Priority Data: 60/071,247 12 January 1998 (12.01.98) US (71) Applicant (for all designated States except US): G.D. SEARLE & CO. [US/US]; Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): VIOLAND, Bernard, N. [US/US]; 1710 Woods Bend Lane, Glencoe, MO 63038 (US). POLAZZI, Joseph, O. [US/US]; Chesterfield, MO (US). CASPERSON, Gerald, F. [US/US]; Ballwin, MO (US). (74) Agents: WILLIAMS, Roger, A. et al.; G.D. Searle & Co., Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>

(54) Title: METHOD OF REFOLDING ANGIOSTATIN



(57) Abstract

An efficient method for expressing, solubilizing, refolding, and purifying, angiostatin from inclusion bodies expressed in bacteria is disclosed. Nucleic acids encoding full-length angiostatin and subfragments encoding kringle domains (K1-4, K2-3, K2-4, K1-3, and K1) are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	JP	Japan	NE	Niger	UZ	Uzbekistan
CG	Congo	KE	Kenya	NL	Netherlands	VN	Viet Nam
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	YU	Yugoslavia
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand	ZW	Zimbabwe
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Method of refolding angiostatin

Priority

The present application claims priority under Title 35, United States Code, § 119 of United States Provisional Application Serial No. 60/071,247 filed January 12, 1998.

Field of the invention

An efficient method for purifying and refolding angiostatin in inclusion bodies expressed in bacteria is disclosed. Nucleic acids encoding full-length angiostatin and subfragments encoding kringle domains K1-4, K2-3, K2-4, K1-3, and K1 are also disclosed.

Background of the invention

Angiogenesis

Angiogenesis, the growth of new blood vessels, plays an important role in cancer growth and metastasis. In humans, the extent of vasculature in a tumor has been shown to correlate with the patient prognosis for a variety of cancers (Folkman, J., *Seminars in Medicine of the Beth Israel Hospital, Boston* 333(26): 1757-1763, 1995; Gasparini, G., *European Journal of Cancer* 32A(14): 2485-2493, 1996; Pluda, J. M., *Seminars in Oncology* 24(2): 203-218, 1997). In normal adults, angiogenesis is limited to well controlled situations, such as wound healing and the female reproductive system (Battegay, E.J., *J Mol Med* 73:333-346, 1995; Dvorak, H.F., *New Engl J Med*, 315: 1650-1659, 1986).

Animal studies suggest that tumors can exist in a dormant state, in which tumor growth is limited by a balance between high rates of proliferation and high rates of apoptosis (Holmgren, L. et al., *Nat. Med. (N. Y.)* 1(2): 149-153, 1995; Hanahan, D. et al., *Cell* 86(3): 353-364, 1996). The switch to an angiogenic phenotype allows tumor cells to escape from dormancy and to grow rapidly, presumably as the result of a decrease in the apoptotic rate of the tumor cells (Bouck, *Cancer Cells*, 2(6): 179-185, 1990; Dameron et al., *Cold Spring Harb Symp Quant Biol*, 59: 483-489, 1994). The control of angiogenesis is thought to be a balance between factors which promote new vessel formation and anti-angiogenic factors which suppress the formation of a

neovasculature (Bouck, N. et al., *Advances in Cancer Research* 69: 135-173, 1996; O'Reilly et al., *Cell (Cambridge, Mass)* 79(2): 315-328, 1994).

5 A variety of pro-angiogenic factors have been characterized including basic and acid fibroblast growth factors (bFGF and aFGF) and vascular permeability factor/vascular endothelial growth factor (VPF/VEGF) (Potgens, A. J. G. et al., *Biol. Chem. Hoppe-Seyler* 376: 57-70, 1995; Ferrara, N., *European Journal of Cancer* 32A(14): 2413-2442, 1996; Bikfalvi, A. et al., *Endocrine Reviews* 18: 26-45, 1997). Several endogenous anti-angiogenic factors have also been characterized, including
10 angiostatin (O'Reilly et al., *Cell* 79(2): 315-328, 1994), endostatin (O'Reilly et al., *Cell* 88(2): 277-285, 1997), interferon-alpha. (Ezekowitz et al., *N. Engl. J. Med.*, May 28, 326(22) 1456-1463, 1992), thrombospondin (Good et al., *Proc Natl Acad Sci USA* 87(17): 6624-6628, 1990; Tolsma et al., *J Cell Biol* 122(2): 497-511, 1993), and platelet factor 4 (PF4) (Maione et al., *Science* 247:(4938): 77-79, 1990).

Angiostatin

15 Angiostatin is a 38 kD protein comprising the first three or four kringle domains of plasminogen and was first described in 1994 (O'Reilly, M. S. et al., *Cell* 79(2): 315-328, 1994). Mice bearing primary (Lewis lung carcinoma-low metastatic) tumors did not respond to angiogenic stimuli such as bFGF in a corneal micropocket assay and the growth of metastatic tumors in these mice was suppressed until the
20 primary tumor was excised. The factor responsible for the inhibition of angiogenesis and tumor growth was designated mouse angiostatin. Angiostatin was also shown to inhibit the growth of endothelial cells in vitro.

Human angiostatin can be prepared by digestion of plasminogen by porcine elastase (O'Reilly et al., *Cell* 79(2): 315-328, 1994) or with human metalloelastase
25 (Dong et al., *Cell* 88, 801-810, 1997). The angiostatin produced via porcine elastase digestion inhibited the growth of metastases and primary tumors in mice. O'Reilly et al. (*Cell* 79(2): 315-328, 1994) demonstrated that human angiostatin inhibited metastasis of Lewis lung carcinoma in SCID mice. The same group (O'Reilly, M. S. et al., *Nat. Med. (N. Y.)* 2(6): 689-692, 1996) subsequently showed that human
30 angiostatin inhibited the growth of the human tumors PC3 prostate carcinoma, clone A colon carcinoma, and MDA-MB breast carcinoma in SCID mice. Human angiostatin also inhibited the growth of the mouse tumors Lewis lung carcinoma, T241 fibrosarcoma and M5076 reticulum cell carcinoma in C57Bl mice. The precise composition of the molecules is not known because these enzymatically-prepared
35 angiostatins are not well characterized biochemically.

Expression of angiostatin

Angiostatins of known composition can be prepared by means of recombinant DNA technology and expression in heterologous cell systems. Recombinant human angiostatin comprising Kringle domains one through four (K1-4) has been produced in the yeast *Pichia pastoris* (Sim et al., *Cancer Res* 57: 1329-1334, 1997). The recombinant human protein inhibited growth of endothelial cells in vitro and inhibited metastasis of Lewis lung carcinoma in C57Bl mice. Recombinant murine angiostatin (K1-4) has been produced in insect cells (Wu et al., *Biochem Biophys Res Comm* 236: 651-654, 1997). The recombinant mouse protein inhibited endothelial cell growth in vitro and growth of primary Lewis lung carcinoma *in vivo*. These experiments demonstrated that the first four kringle domains are sufficient for angiostatin activity but did not determine which kringle domains are necessary.

Cao et al. (*J. Biol. Chem.* 271: 29461-29467, 1996), produced fragments of human plasminogen by proteolysis and by expression of recombinant proteins in *E. coli*. These authors showed that kringle one and to a lesser extent kringles two and three of plasminogen were responsible for the inhibition of endothelial cell growth in vitro. Specifically, kringles 1-4 and 1-3 inhibited at similar concentrations, while K1 alone inhibited endothelial cell growth at four-fold higher concentrations. Kringles two and three inhibited to a lesser extent. More recently Cao et al. (*J Biol Chem* 272: 22924-22928, 1997), showed that recombinant mouse or human kringle five inhibited endothelial cell growth at lower concentrations than angiostatin (K1-4). These experiments demonstrated in vitro angiostatin-like activity but did not address in vivo action against tumors and their metastases.

World patent applications WO 95/29242 A1, WO 96/41194 A1, WO 96/35774 A2, and WO 98/54217 A1 describe the expression, purification, and characterization of angiostatin. WO 95/29242 A1 951102 discloses purification of a protein from blood and urine by HPLC that inhibits proliferation of endothelial cells. The protein has a molecular weight between 38 kilodaltons and 45 kilodaltons and an amino acid sequence substantially similar to that of a murine plasminogen fragment beginning at amino acid number 79 of a murine plasminogen molecule. WO 96/41194 A1 961219, discloses compounds and methods for the diagnosis and monitoring of angiogenesis-dependent diseases. WO 96/35774 A2 961114 discloses the structure of protein fragments, generally corresponding to kringle structures occurring within angiostatin. It also discloses aggregate forms of angiostatin, which have endothelial cell inhibiting activity, and provides a means for inhibiting angiogenesis of tumors and for treating

angiogenic-mediated diseases. WO 98/54217 A1 describes fragments of angiostatin that have endothelial cell inhibiting activity.

Anti-angiogenic therapy may offer several advantages over conventional chemotherapy for the treatment of cancer. Anti-angiogenic agents have low toxicity in preclinical trials and development of drug resistance has not been observed (Folkman, J., *Seminars in Medicine of the Beth Israel Hospital, Boston* 333(26): 1757-1763, 1995). As angiogenesis is a complex process, made up of many steps including invasion, proliferation and migration of endothelial cells, it can be anticipated that combination therapies may be most effective. In fact, combinations of chemotherapy with anti-angiogenic therapy have already shown promising results in pre-clinical models (Teicher, B. A. et al., *Breast Cancer Research and Treatment* 36: 227-233, 1995; Teicher, B. A. et al., *European Journal of Cancer* 32A(14): 2461-2466, 1996).

Purification and refolding of angiostatin

Although many types of expression systems have been developed over the past twenty years, bacterial systems, particularly those based on *E. coli*, are widely used for the production of proteins on an industrial scale. Vectors which permit high level expression and the ability to carry out fermentations at high cell densities and low cost, have contributed to the extensive development and use of *E. coli*-based expression systems. One significant problem, however, is the tendency of *E. coli* to form inclusion bodies which contain the desired recombinant protein. Inclusion body formation necessitates additional downstream processing, such as *in vitro* refolding, before biologically active proteins can be recovered. The tendency to form insoluble aggregates does not appear to correlate with factors such as size, hydrophobicity, subunit structure, or the use of fusion domains (Kane J.F. and Harley, D.L., *Tibtech* 6: 95, 1988). Inclusion body formation also appears to be determined by the rates of protein synthesis, folding, aggregation, and proteolytic degradation, the solubility and thermodynamics of folding intermediates and native proteins, and the interactions of these species with chaperone proteins (Rainer Rudolph, *In Protein Engineering: Principles and Practice*, Edited by Jeffrey L. Cleland and Charles S. Craik, p 283-298, Wiley-Liss, Inc., New York, New York, 1996).

Inclusion bodies generally form in the cytoplasm of cells expressing a recombinant protein at high levels. They refract light when observed by phase contrast microscopy and thus are sometimes referred to as refractile bodies. The inclusion bodies are characterized by a relatively high specific density and can be pelleted from lysed cells by centrifugation. The formation of inclusion bodies may

protect recombinant proteins from proteolysis as they do not easily disintegrate under physiological solvent conditions. High concentrations of denaturants, such as 6 M guanidine hydrochloride or 6-8 M urea, have been commonly used to solubilize the proteins present in inclusion bodies. A variety of inclusion body solubilization protocols have been compared (Fisher, B., Summer, L. and Goodenough, P. *Biotechnol. Bioeng.* 1: 3-13, 1992).

Although the desired foreign gene product is the main component of inclusion bodies, other host cell proteins such as small heat shock proteins, outer membrane proteins, elongation factor EF-Tu, and RNA polymerase may also be enriched in such preparations (Allen, S.P., Polazzi, J.O. Gierse, J.K., and Easton, A.M., *J. Bacteriol.* 174: 6938-6947, 1992); Hart, R.A., Rinas, U., and Bailey, J.E., *J. Biol. Chem.* 265: 12728-12733, 1990; Hartley, D.L., and Kane, J.F., *Biochem. Soc. Trans.* 16: 101, 1988).

Detailed studies of the refolding of plasminogen-derived kringle domains isolated from bacterial inclusion bodies all suggest that optimal refolding occurs at pH 8.0 (Trexler, M., and Patty, L., *Proc. Natl. Acad. Sci.* 80: 2457-2461, 1983; Menhart, N., et al., *Biochemistry* 30: 1948-1957, 1991; Marti, D., et al., *Eur. J. Biochem.*, 219: 455-462, 1994; Cao, Y., et al., *J. Biol. Chem.* 271: 29461-29467, 1996). In these studies, inclusion bodies are generally dissolved in high concentrations of a denaturant (such as urea or guanidine hydrochloride) in the presence of a reducing agent (such as dithiothreitol or beta-mercaptoethanol). In each of these references, the denaturant is diluted either rapidly or slowly to initiate refolding.

World patent application WO 96/35,774 to Folkman et al., filed April 26, 1996, describes proteins derived from human plasminogen which are capable of inhibiting endothelial cell proliferation. Methods for the preparation of aggregate forms of these proteins are also disclosed.

Summary of the Invention

One object of the invention is to describe a method for expressing high levels of angiostatin and other plasminogen kringles in bacteria.

Another aspect of the invention is to describe an efficient method by which angiostatin inclusion bodies can be solubilized, subsequently refolded, and purified to generate biologically active material.

Preferably the steps of solubilizing and refolding angiostatin inclusion bodies are carried out at an elevated pH, ranging from about pH 9 to about pH 11.5.

Preferably the pH range is from about pH 10 to about pH 11. Even more preferably, the pH is 10.5.

5 Preferably, the concentration of angiostatin gene product is present at a concentration of from about 1 to about 20 mg/ml during the solubilization step. Even more preferably, the concentration is about 6 mg/ml.

Preferably, the concentration of angiostatin gene product is present at a concentration of from about 0.1 to about 5 mg/ml during the refolding step. Even more preferably, the concentration is about 1 mg/ml.

10 Preferably a denaturant selected from urea and guanidine hydrochloride is used during the solubilization and refolding steps. Even more preferably, the denaturant is urea.

Preferably the concentration of urea is from about 4 M to about 10 M during the solubilization step. Even more preferably, the concentration is about 6 M.

15 Preferably the concentration of urea is from about 0.5 M to about 3 M during the refolding step. Even more preferably, the concentration is about 1 M.

Preferably the concentration of guanidine hydrochloride is from about 3 M to about 8 M during the solubilization step. Even more preferably, the concentration is about 5 M.

20 Preferably the concentration of guanidine hydrochloride is from about 0.2 M to about 2 M during the refolding step. Even more preferably, the concentration is about 0.5 M.

25 Preferably, the solubilization and reducing steps are carried out in the presence of a reducing agent capable of reducing disulfide linkages to sulfhydryl groups. Preferably the reducing agent is selected from the group consisting of DTT, BME, cysteine, and reduced glutathione. Even more preferably, the reducing agent is DTT.

Preferably, DTT is present at a concentration of from about 2 mM to about 10 mM during the solubilization step. Even more preferably, the concentration is about 6 mM.

30 Preferably, DTT is present at a concentration of from about 0.5 mM to about 2 mM during the refolding step. Even more preferably, the concentration is about 1 mM.

Preferably, reduced glutathione is present at a concentration of from about 5 mM to about 20 mM during the solubilization step. Even more preferably, the concentration is about 10 mM.

5 Preferably, reduced glutathione is present at a concentration of from about 1 mM to about 4 mM during the refolding step. Even more preferably, the concentration is about 2 mM.

Preferably, cysteine is present at a concentration of from about 5 mM to about 20 mM during the solubilization step. Even more preferably, the concentration is about 10 mM.

10 Preferably, cysteine is present at a concentration of from about 1 mM to about 4 mM during the refolding step. Even more preferably, the concentration is about 2 mM.

15 Preferably, an agent capable of enhancing the interchange of disulfide bonds is present during the refolding step. Preferably the agent is selected from cystine and oxidized glutathione. Even more preferably, the agent is cystine.

Preferably, cystine is present at a concentration of from about 0.2 mM to about 5 mM during the refolding step. Even more preferably, the concentration is about 1 mM.

20 Preferably the refolding step is carried out in the presence of a ligand capable of binding kringle domains of angiostatin and plasminogen. Preferably the ligand is selected from the group consisting of 6-amino hexanoic acid, 7-amino heptanoic acid, 5-amino pentanoic acid, and 4-amino butyric acid. Even more preferably, the ligand is 6-amino hexanoic acid.

25 Preferably disulfide bonds are formed through air oxidation during the refolding step. Preferably, the air oxidation step is carried out from about 12 to about 96 hours. Even more preferably, the air oxidation step is carried out from about 24 to about 72 hours. Most preferably the air oxidation step is carried out about 60 hours.

30 Preferably refolded angiostatin is further purified by a process selected from the group consisting of lysine-sepharose chromatography, ion-exchange chromatography, hydrophobic interaction chromatography and RP-HPLC.

Preferably the method of expression, solubilization, refolding and purification uses an plasminogen-derived kringle domains or angiostatin genes that are human.

Even more preferably these genes are selected from the group consisting of SEQ ID NO: 10 through SEQ ID NO: 29.

5 Preferably the method of expression, solubilization, refolding and purification uses the products of plasminogen-derived kringle domains or angiostatin genes that are human. Even more preferably the products of these genes are encoded by a sequence selected from the group consisting of SEQ ID NO: 10 through SEQ ID NO: 29.

10 Another aspect of the invention is to describe novel polypeptides derived from human plasminogen. Preferably these polypeptides are encoded by a sequence selected from the group consisting of SEQ ID NO: 10 through SEQ ID NO: 29.

Novel proteins of this invention are a modified human angiostatin amino acid sequence, and said protein can optionally be immediately preceded by (methionine⁻¹), (alanine⁻¹), (methionine⁻², alanine⁻¹), (serine⁻¹), (methionine⁻², serine⁻¹), (cysteine⁻¹), or (methionine⁻², cysteine⁻¹).

15 Preferably these proteins are selected from the group consisting of angiostatin K1, angiostatin K5, angiostatin K1-K3, angiostatin K1-K4, angiostatin K1-K5.

20 Additionally, the present invention relates to recombinant expression vectors comprising nucleotide sequences encoding the angiostatin, variants and muteins of angiostatin, related microbial and eukaryotic expression systems, and processes for making (comprising the steps of expressing, solubilizing, refolding, purifying) these proteins.

25 Cloning of DNA sequences encoding these proteins may be accomplished by the use of intermediate vectors. Alternatively, one gene can be cloned directly into a vector containing the other gene. Linkers and adapters can be used for joining the DNA sequences, as well as replacing lost sequences, where a restriction site was internal to the region of interest. Thus genetic material (DNA) encoding one polypeptide, peptide linker, and the other polypeptide is inserted into a suitable expression vector which is used to transform bacteria, yeast, insect cells or mammalian cells. The transformed organism or cell line is grown and the protein isolated by standard techniques. The resulting product is therefore a new protein which has all or a portion of one protein joined by a linker region to a all or a portion of second protein.

30

Another aspect of the present invention includes plasmid DNA vectors for use in the expression of these proteins. These vectors contain the novel DNA sequences

described above which code for the novel polypeptides of the invention. Appropriate vectors which can transform microorganisms or cell lines capable of expressing the proteins include expression vectors comprising nucleotide sequences coding for the proteins joined to transcriptional and translational regulatory sequences which are selected according to the host cells used.

Vectors incorporating modified sequences as described above are included in the present invention and are useful in the production of the proteins. The vector employed in the method also contains selected regulatory sequences in operative association with the DNA coding sequences of the invention and which are capable of directing the replication and expression thereof in selected host cells.

Methods for producing these proteins is another aspect of the present invention. The method of the present invention involves culturing suitable cells or cell lines, which has been transformed with a vector containing a DNA sequence coding for expression of a novel multi-functional protein. Suitable cells or cell lines may be bacterial cells. For example, the various strains of *E. coli* are well-known as host cells in the field of biotechnology. Examples of such strains include *E. coli* strains JM101 (Yanisch-Perron et al., *Gene* 33: 103-119, 1985) and MON105 (Obukowicz et al., *Applied Environmental Microbiology* 58: 1511-1523, 1992). Also included in the present invention is the expression of the multi-functional proteins utilizing a chromosomal expression vector for *E. coli* based on the bacteriophage Mu (Weinberg et al., *Gene* 126: 25-33, 1993). Various strains of *B. subtilis* may also be employed in this method. Many strains of yeast cells known to those skilled in the art are also available as host cells for expression of the polypeptides of the present invention.

When expressed in the *E. coli* cytoplasm, the gene encoding the proteins of the present invention may also be constructed such that at the 5' end of the gene codons are added to encode Met²-Ala⁻¹, Met²-Ser⁻¹, Met²-Cys⁻¹, or Met⁻¹ at the N-terminus of the protein. The N termini of proteins made in the cytoplasm of *E. coli* are affected by post-translational processing by methionine aminopeptidase (Ben Bassat et al., *J. Bacteriol.* 169:751-757, 1987) and possibly by other peptidases so that upon expression the methionine is cleaved off the N-terminus. The proteins of the present invention may include polypeptides having Met⁻¹, Ala⁻¹, Ser⁻¹, Cys⁻¹, Met²-Ala⁻¹, Met²-Ser⁻¹, or Met²-Cys⁻¹ at the N-terminus. These mutant proteins may also be expressed in *E. coli* by fusing a secretion signal peptide to the N-terminus. This signal peptide is cleaved from the polypeptide as part of the secretion process.

Definitions

The following is a list of abbreviations and the corresponding meanings as used interchangeably herein:

5 BME = beta mercaptoethanol
DTT = dithiothreitol
g = gram(s)
HPLC = high performance liquid chromatography
mg = milligram(s)
ml or mL = milliliter(s)
10 RT = room temperature
ug or µg = microliter(s)
ul or µl or uL or µL = microliter(s)

The following is a list of definitions of various terms used herein:

15 The term "anti-tumor" means possessing an activity which slows or abolishes the growth of, or which kills, or otherwise harms tumors *in vivo*.

The term "native sequence" refers to an amino acid or nucleic acid sequence which is identical to a wild-type or native form of a gene or protein.

20 The terms "mutant amino acid sequence," "mutant protein", "variant protein", "mutein", or "mutant polypeptide" refer to a polypeptide having an amino acid sequence which varies from a native sequence due to amino acid additions, deletions, substitutions, or any combination thereof, or is encoded by a nucleotide sequence from an intentionally-made variant derived from a native sequence or chemically synthesized..

25 The term "angiostatin" means a protein fragment of plasminogen (encoded by SEQ ID NO: 9) having anti-angiogenic activity. The activity of said fragments can be determined by the mouse corneal micropocket assay of angiogenesis or by inhibition of endothelial cell growth *in vitro*.

30 The term "kringle" or "K" means a fragment or substructure of plasminogen or angiostatin bounded by disulfide-linked cysteine residues, and which retain a structure similar to that which it will adopt within the whole of plasminogen. A kringle domain may also include interkringle sequences which serve to connect it with other kringle domains. The terms K1, K2, K3, K4, K5 refer to specific kringle domains.

and terms like K1-3 or K2-4, for example, refer to contiguous protein segments from the first kringle through the last kringle in a series, including the interkringle sequences

Brief description of the figures

5 **Figure 1 shows a schematic of selected cloned angiostatin fragments containing kringle domains**

10 The kringle domains of human plasminogen pMON24324 (K1-5), pMON24642 (K1-4E), pMON24643 (K1-4), pMON24644 (2-3), pMON24645 (K2-4), pMON24646 (K1-3), pMON24649 (K1) are shown. The labels K1, K2, K3, K4, and K5, refer to
15 specific kringle domains within human plasminogen or angiostatin. Plasmid pMON24624 contains the sequence of human plasminogen, K1-K5, including the N-terminal peptide (NTP) and the signal peptide (SP). The plasmid pMON24642 encodes kringles K1-K4 plus additional human plasminogen sequences at its 5' and 3' ends shown in the diagram by solid areas. Schematic depictions of the angiostatin-related sequences in plasmids pMON24625 (K1-K4), pMON24626 (K1-K3),
20 pMON24627 (K2-K4), pMON24628 (K2-K3), and pMON24650 (K1) which contain specific kringle domains of human plasminogen are also shown. These plasmids also contain sequences necessary for the expression of the encoded polypeptides in *E. coli*.

Figure 2 shows a schematic of the angiostatin refolding procedure

20 Optimal solubilization conditions in the presence of urea and DTT and refolding in the presence of cystine and 6-amino hexanoic acid are outlined.

Figure 3 shows the inhibition of HMEC migration by angiostatin

25 Purified K1-3, K1-4, and elastase-derived angiostatin were assayed in the HMEC cell migration assay at 20 ug/ml. All angiostatins showed inhibition of migration.

Figure 4 shows the inhibition of BAEC migration by angiostatin

Purified K1-3, K1-4, and elastase-derived angiostatin were assayed in the BAEC cell migration assay at 20 ug/ml. All angiostatins showed inhibition of migration.

Figure 5 shows the inhibition of HMEC migration by angiostatin

Inhibition of HMEC cell migration versus angiostatin concentration. Dose response curves of angiostatin concentration versus % inhibition of cell migration are shown for K1-3, K1-4, and elastase-derived angiostatin.

5 **Figure 6 shows binding of refolded angiostatin to lysine-sepharose columns at pH 8.0 and 10.5**

Elution profiles for K1-3 at pH 10.5 and single kringle at pH 8.0 by HPLC lysine chromatography are shown.

Detailed description of the invention

10 The following examples will illustrate the invention in greater detail, although it will be understood that the invention is not limited to these specific examples. Various other examples will be apparent to the person skilled in the art after reading the present disclosure without departing from the spirit and scope of the invention. It is intended that all such other examples be included within the scope of the appended
15 claims.

General methods

General methods of cloning, expressing, and characterizing proteins are found in T. Maniatis et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982, and references cited therein, incorporated herein by reference; and
20 in J. Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory, 1989, and references cited therein, incorporated herein by reference.

Unless noted otherwise, all specialty chemicals were obtained from Sigma, Co. (St. Louis, MO). Restriction endonucleases and T4 DNA ligase were obtained from
25 New England Biolabs (Beverly, MA) or Boehringer Mannheim (Indianapolis, IN).

Strains and plasmids

The bacterial strains used in these studies are listed in Table 1. Plasmids used in or constructed for this study are listed in Table 2.

Transformation of *E. coli* strains

E. coli strains (Table 1), such as DH10B™ (Life Technologies, Gaithersburg, MD) and TG1 (Amersham Corp., Arlington Heights, IL) are used for transformation of ligation reactions and are the hosts used to prepare plasmid DNA for transfecting mammalian cells. *E. coli* strains, such as JM101 (Yanisch-Perron et al., *Gene*, 33: 103-119, 1985) and MON105 (Obukowicz et al., *Appl. and Envir. Micro.*, 58: 1511-1523, 1992) can be used for expressing the proteins of the present invention in the cytoplasm or periplasmic space.

DH10B™ Subcloning efficiency cells are purchased as competent cells and are ready for transformation using the manufacturer's protocol, while both *E. coli* strains JM101 AND MON105 are rendered competent to take up DNA using a CaCl₂ method. Typically, 20 to 50 mL of cells are grown in LB medium (1% Bacto-tryptone, 0.5% Bacto-yeast extract, 150 mM NaCl) to a density of approximately 1.0 absorbance unit at 600 nanometers (OD600) as measured by a Baush & Lomb Spectronic spectrophotometer (Rochester, NY). The cells are collected by centrifugation and resuspended in one-fifth culture volume of CaCl₂ solution [50 mM CaCl₂, 10 mM Tris-Cl ((10 mM 2-amino-2-(hydroxymethyl) 1,3-propanediol hydrochloride, pH7.4] and are held at 4°C for 30 minutes. The cells are again collected by centrifugation and resuspended in one-tenth culture volume of CaCl₂ solution. Ligated DNA is added to 0.2 mL of these cells, and the samples are held at 4°C for 30-60 minutes. The samples are shifted to 42°C for two minutes and 1.0 mL of LB is added prior to shaking the samples at 37°C for one hour. Cells from these samples are spread on plates (LB medium plus 1.5% Bacto-agar) containing either ampicillin (100 micrograms/mL, ug/mL) when selecting for ampicillin-resistant transformants, or spectinomycin (75 ug/mL) when selecting for spectinomycin-resistant transformants. The plates are incubated overnight at 37°C.

Colonies are picked and inoculated into LB plus appropriate antibiotic (100 ug/mL. ar .picillin or 75 ug/mL spectinomycin) and are grown at 37°C while shaking.

DNA isolation and characterization

Plasmid DNA can be isolated by a number of different methods and using commercially available kits known to those skilled in the art. Plasmid DNA is isolated using the Promega Wizard™ Miniprep kit (Madison, WI), the Qiagen QIAwell Plasmid isolation kits (Chatsworth, CA) or Qiagen Plasmid Midi or Mini kit. These kits follow the same general procedure for plasmid DNA isolation. Briefly, cells are pelleted by

centrifugation (5000 x g), the plasmid DNA released with sequential NaOH/acid treatment, and cellular debris is removed by centrifugation (10000 x g). The supernatant (containing the plasmid DNA) is loaded onto a column containing a DNA-binding resin, the column is washed, and plasmid DNA eluted. After screening for the colonies with the plasmid of interest, the *E. coli* cells are inoculated into 50-100 ml of LB plus appropriate antibiotic for overnight growth at 37°C in an air incubator while shaking. The purified plasmid DNA is used for DNA sequencing, further restriction enzyme digestion, additional subcloning of DNA fragments and transfection into *E. coli*, mammalian cells, or other cell types.

Sequence confirmation

Purified plasmid DNA is resuspended in deionized H₂O and its concentration is determined by measuring the absorbance at 260/280 nm in a Bausch and Lomb Spectronic 601 UV spectrometer. DNA samples are sequenced using ABI PRISM™ DyeDeoxy™ terminator sequencing chemistry (Applied Biosystems Division of Perkin Elmer Corporation, Lincoln City, CA) kits (Part Number 401388 or 402078) according to the manufacturer's suggested protocol usually modified by the addition of 5% DMSO to the sequencing mixture. Sequencing reactions are performed in a DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT) following the recommended amplification conditions. Samples are purified to remove excess dye terminators with Centri-Sep™ spin columns (Princeton Separations, Adelphia, NJ) and lyophilized. Fluorescent dye labeled sequencing reactions are resuspended in deionized formamide, and sequenced on denaturing 4.75% polyacrylamide-8M urea gels using ABI Model 373A and Model 377 automated DNA sequencers. Overlapping DNA sequence fragments are analyzed and assembled into master DNA contigs using Sequencher DNA analysis software (Gene Codes Corporation, Ann Arbor, MI).

Small-scale expression of angiostatin in *E. coli*

E. coli strain MON105 or JM101 harboring the plasmid of interest is grown at 37°C in M9 plus casamino acids medium with shaking in an air incubator Model G25 from New Brunswick Scientific (Edison, NJ). Growth is monitored at OD₆₀₀ until it reaches a value of 1.0 at which time nalidixic acid (10 mg/mL) in 0.1 N NaOH is added to a final concentration of 50 µg/mL. The cultures are then shaken at 37°C for three to four additional hours. A high degree of aeration is maintained throughout THE culture period in order to achieve maximal production of the desired gene product. The cells are examined under a light microscope for the presence of inclusion bodies (IB). One mL aliquots of the culture are removed for analysis of protein content by

boiling the pelleted cells, treating them with reducing buffer and electrophoresis via SDS-PAGE (see Maniatis et al., 'Molecular Cloning: A Laboratory Manual', 1982). The culture is centrifuged (5000 x g) to pellet the cells.

Large (10L)-scale expression of angiostatin in *E. coli*

5 The Angiostatin molecules were all scaled up in a 10 L Biostat E™ fermenter (B. Braun Biotech Inc., Allentown PA). The fermentation medium used was M9 salts supplemented with 2% casamino acids (Difco Laboratories, Detroit MI) and glucose. Approximately one milliliter of a thawed culture was transferred to a 3.8 L Fernbach
10 shake flask containing 1.0 L of medium and then incubated at 37°C at a shaker speed of 250 rpm for 12 hours. This shake flask culture was then used to inoculate a 10 L fermenter containing 9.0 L of medium. The fermentation conditions were as follows: agitation = 1000 rpm, sparger airflow rate = 15 liters/min, pH control = 7.0 by the addition of NH₄OH, backpressure = 10 psi, dissolved oxygen control > 30%, and
15 temperature = 37°C. Glucose was initially batched in at 15 g/l and controlled at 2-5 g/l by the addition of a 50% glucose feed stock. The fermentation culture was grown to an initial OD (550nm) = 12-15 and induced with 50 mg/l nalidixic acid. The fermentation was harvested four hours post induction by continuous-flow centrifugation.

10L scale inclusion body isolation

20 The cell paste from a 10L fermentation was resuspended in approximately 6.0 L of 50 mM Tris/150 mM EDTA buffer containing 40 pellets of 50 ml Complete Protease Inhibitor Cocktail (Boehringer Mannheim; Indianapolis, IN) and 600,000 KIU of Trasylol. The resuspension was passed once through a Microfluidizer (Newton, MA) at 10,000 psi and temperature maintained under 8°C. The recovered homogenate was then spun at 15,000 x G for 25 minutes and mixed to consistency with an Ultr-
25 Turrax mixer. The cell paste resuspension was then homogenized by passing the cell suspension two times through a Microfluidizer (Newton, MA) operating at a pressure of 10,000 psi. The temperature of the homogenate was maintained less than 6°C by collection in a stainless steel container placed in an ice bath. Inclusion bodies were then isolated by centrifugation of the cell homogenate at 15,000 X g for 30 minutes
30 and removal of supernatant. Inclusion bodies were then washed a total of two times by resuspending the inclusion bodies in chilled D.I. and centrifugation at 15,000 X g for 30 minutes. Angiostatin Inclusion bodies were then frozen at -70°C.

Small-scale isolation of inclusion bodies

The cell pellet from a 330 mL *E. coli* culture is resuspended in 15 mL of sonication buffer Tris-Cl, pH 8.0 + 1 mM ethylenediaminetetraacetic acid (EDTA). These resuspended cells are sonicated using the microtip probe of a Sonicator Cell Disruptor (Model W-375, Heat Systems-Ultrasonics, Inc., Farmingdale, New York). Three rounds of sonication in sonication buffer followed by centrifugation are employed to disrupt the cells and wash the inclusion bodies (IB). The first round of sonication is a 3 minute burst followed by a 1 minute burst, and the final two rounds of sonication are for 1 minute each.

Extraction and refolding of proteins from inclusion body pellets

All steps are at 4°C. Angiostatin K1-4 inclusion bodies were dissolved in 6M urea, 5 mM DTT, 50 mM Bis-Tris Propane, pH 10.8 at 1.2 mg/ml. This solution was stirred for 2 hours and then 6-amino hexanoic acid (stock 1M) was added to achieve a 6 mM concentration, followed by addition of cystine (stock 0.2M, pH 10.5) to 6 mM. This was mixed for 5 minutes and then the K1-4 angiostatin was diluted to 0.2 mg/ml by addition of 75 mM Bis-Tris propane, pH 10.5. It was then stirred for 72 hours to complete the refolding of the protein as assessed by reverse-phase HPLC or by binding to Lys-HPLC (Figure 6).

Purification

Purification of angiostatin K1-4 was achieved using a Sepharose-Lys affinity column followed by Q-Sepharose chromatography. The sample was next dialyzed against 50 mM sodium phosphate, pH 8.0 and then filtered through a 1 µm filter to clarify. This sample was loaded onto a Sepharose-Lys column equilibrated in 50 mM sodium phosphate, pH 8.0. The load was at approximately 1.5 gm of K1-4 per ml of resin. After loading the column, it was washed with 1 column volume of equilibrating buffer and then K1-4 was eluted using a 10 column volume gradient from 0-8 mM 6-amino hexanoic acid in 50 mM sodium phosphate, pH 8.0.

Fractions were analyzed by SDS-gel electrophoresis and/or RP-HPLC and pooled. This pool was made 40 mM Tris-Cl, and adjusted to pH 9.0. It was then dialyzed extensively vs. 40 mM Tris-Cl, pH 9.0 and applied to a Pharmacia Q-Sepharose HP column (~5 mgs protein/ml resin) equilibrated in the same buffer. K1-4 was then eluted using a NaCl gradient from 0-0.1M in the equilibrating buffer. Fractions were analyzed by RP-HPLC and/or SDS gel electrophoresis and pooled.

In some cases the folded proteins can be affinity-purified using affinity reagents such as monoclonal antibodies or receptor subunits attached to a suitable matrix. Purification can also be accomplished using any of a variety of chromatographic methods such as: ion exchange, gel filtration or hydrophobic interaction chromatography or reversed phase HPLC. These and other protein purification methods are described in detail in *Methods in Enzymology, Volume 182 "Guide to Protein Purification"* edited by Murray Deutscher, Academic Press, San Diego, California, 1990.

Protein characterization

The purified protein is analyzed by RP-HPLC, electrospray mass spectrometry, and SDS-PAGE. The protein quantitation is done by amino acid composition, RP-HPLC, and Bradford protein determination. In some cases tryptic peptide mapping is performed in conjunction with electrospray mass spectrometry to confirm the identity of the protein.

Inhibitory effect on proliferation of Daudi cells

Human Burkitt's lymphoma Daudi cells (ATCC) are seeded at 2×10^4 cells/well into 96-well tissue culture plates and cells are cultured in the presence or absence of serial doses of rhIFN.alpha. 2b for 3 days. Cultures are pulsed with ^3H -thymidine for the last hour of the culture period, and the ^3H -thymidine uptake, counts per minute (cpm), ARE measured on a Beta-plate reader. All samples are assayed in triplicate.

Endothelial cell proliferation assay

The endothelial cell proliferation assay was performed as described by Cao et al. (*J. Biol. Chem.* 271: 29461-29467, 1996). Briefly, human dermal microvascular endothelial cells (HdMVEC, Clonetics) or bovine microvascular endothelial cells (BacEnd, Incell) were maintained in MCDB131 containing 5% heat-inactivated fetal bovine serum (FBS, Hyclone), antibiotics, 100 ug/ml heparin (Sigma) and 100 ug/ml endothelial mitogen (Biomedical Technologies). Confluent monolayers at passages 2-5 were dispersed in 0.05% trypsin and resuspended in complete medium. Five hundred ul of complete media containing 1.25×10^4 cells were seeded into wells of a 24-well tissue culture plate coated with 0.1% gelatin (Sigma). The cells were incubated overnight at $37^\circ\text{C}/5\% \text{CO}_2$ at which time the media was replaced with 250 ul of media containing 5% FBS and various concentrations of inhibitors. After 30 minutes of incubation, 250 ul of media containing 1 ng/ml bFGF (R&D Systems) was added and

the cells were incubated for an additional 72 hours, at which time they were trypsinized and counted with a Coulter counter.

Endothelial cell *in vitro* tube formation assay

The tube formation assay is performed as previously described (Gately et al., *Cancer Res.* 56:4887-4890, 1996). The effect of angiogenesis inhibitors can be tested in an *in vitro* angiogenesis assay; endothelial cell tube formation on matrix. Matrigel (Becton Dickinson) was thawed on ice and diluted 1:1 with MCDB 131 medium supplemented with 2 mM L-glutamine, containing either buffer control or testing compounds. Aliquots of 0.5 ml of 1:1 diluted matrigel were plated into individual wells of 2-chambered Tissue-Tek polystyrene slides (Nunc) and allowed to polymerize at 37°C for at least 30 min. Primary bovine capillary endothelial cells derived from the bovine adrenal cortex (purchased from Incell) were cultured (as described for bovine endothelial cell proliferation). Following trypsinization, the cells were washed three times, then resuspended at 3.5×10^5 cells/ml in MCDB131 medium supplemented with L-glutamine and 1% fetal bovine serum containing either the same buffer control or testing compounds. The endothelial cells (0.5 ml at 3.5×10^5 cells /ml) were plated onto matrigel coated wells containing the same additives. After incubation for 16-24 hours at 37°C in a 5% CO₂ incubator, cultures were evaluated for tube formation (sprouting). Random fields of each experimental condition were photographed using a Nikon N6006 camera under a Nikon Diaphot microscope. The photographs were then quantitated by measuring the total length of the tubes. Tube formation can also be quantitated using image analysis software. Endothelial cell tube networks were fixed in 100% methanol at -20°C for 7-10 min, rinsed 4 times with phosphate buffered saline (PBS) and incubated overnight at 4°C with rabbit polyclonal anti-human Factor VIII-related antigen (von Willebrand factor, vWF) (Dako). The next day, the cells were stained with a secondary antibody, Cy3-conjugated goat anti-rabbit IgG F(ab')₂ specific (Jackson ImmunoResearch Laboratories). Tube formation was visualized using a Nikon microphot-FXA with a fluorescence filter linked to a computer containing image analysis software. Tube formation was measured as follows: Total area encompassed by endothelial cell tubes/Total endothelial cell surface area.

Endothelial cell migration assay

The endothelial cell migration assay is performed essentially as previously described (Gately et al., *Cancer Res.* 56:4887-4890, 1996). To determine the ability of angiostatin to inhibit the migration of endothelial cells, migration assays were performed in a transwell chamber (Costar) containing 8 mm pore size polycarbonate

membranes. The cells utilized in the assay were either human microvascular endothelial cells from Emory or bovine endothelial cells (kindly provided by Gately Northwestern University, Evanston, IL). The cells were starved overnight before use in MCDB131 + 0.1% BSA (human cells) or DMEM+0.1%BSA (bovine cells), harvested, and resuspended in the same media at 10^6 cells/ml. The lower side of the transwells were coated with 0.1% gelatin for 30 minutes at 37°C before addition of 2×10^5 cells to the upper chamber. The transwell was moved to a well containing the chemoattractant (bFGF or VEGF) in the lower chamber. Migration was allowed to occur overnight at 37°C. The membranes were then fixed and stained, and the number of cells that migrated to the lower side of the membrane counted in 3 high powered fields.

Example 1: Construction of pMON24642 and selection of strains producing high levels of angiostatin K1-4E

A cDNA encoding angiostatin K1-4E (see e.g., pMON24642-a12.seq, SEQID # 10) was synthesized by polymerase chain reaction using the oligonucleotide primers angeg-s1 (anged-s1.seq, SEQID # 1) and angeg-n1 (anged-n1.seq, SEQID # 4) and using the human plasminogen cDNA (pMON24624.seq SEQID # 9) as template. The resulting DNA was cleaved with restriction endonucleases *Afl*III and *Hind*III and was ligated into the *E. coli* expression vector pMON5723 (Clins and Rangwala, *Methods Enzymol.* 185 (Gene Expression Technol.): 115-119, 1990) which had been digested with *Nco*I and *Hind*III. A portion of the ligation reaction was used to transform *E. coli* strain DH10B to spectinomycin resistance.

Plasmid DNA was isolated from the DH10B bacterial colonies and used to transform *E. coli* strain MON105 to spectinomycin resistance. Individual antibiotic resistant colonies were inoculated into 1 ml of Luria Broth with 75 ug/ml spectinomycin and grown over night at 37 degrees with shaking. Fifty microliters of the overnight culture were inoculated into one ml of M9 medium with casamino acids and grown for four hours at 37 degrees with shaking. Nalidixic acid (50 ul of a 1 mg/ml aqueous solution) was added and the cultures were incubated at 37 degrees for an additional three hours. Cultures were examined by phase-contrast microscopy for the presence of cytoplasmic inclusion bodies and by SDS polyacrylamide gel electrophoresis for the presence of the expressed protein. Three cultures expressing the most angiostatin were selected for further characterization. Plasmid DNA was isolated from clones A12, B5, and B6 and was analyzed by DNA sequencing. The sequences of the cDNAs in these plasmids are shown in 24642-A12.seq, 24642-B5.seq and 24642-B6.seq (SEQID # 10, SEQID # 11, and SEQID # 12, respectively). Each

cDNA encoded the expected protein but contained various silent mutations at the third positions of codons two through eight. These mutations were caused by the redundancies in the 5' PCR primer (angec-s1.seq, SEQID # 1), and the particular mutations in these clones result in high level expression of angiostatin. Each cDNA differed from the others and from the wild-type sequence at one or more of these positions.

Example 2: Construction of pMON24643 and selection of strains producing high levels of angiostatin K1-4

A cDNA encoding angiostatin K1-4 was synthesized by polymerase chain reaction using the oligonucleotide primers angec-s2 (angec-s2.seq, SEQID # 2) and angec-n2 (angec-n2.seq, SEQID # 5) and using the human plasminogen cDNA (pMON24624.seq, SEQID # 9) as template. The resulting cDNA was cleaved with restriction endonucleases *AflIII* and *HindIII* and was ligated into the *E. coli* expression vector pMON5723 which had been digested with *NcoI* and *HindIII*. A portion of the ligation reaction was used to transform *E. coli* strain DH10B to spectinomycin resistance

Plasmid DNA was isolated from the DH10B bacterial colonies and used to transform *E. coli* strain MON105 to spectinomycin resistance. Individual antibiotic resistant colonies were inoculated into 1 ml of Luria Broth with 75 ug/ml spectinomycin and grown over night at 37 degrees with shaking. Fifty microliters of the overnight culture were inoculated into one ml of M9 medium with casamino acids and grown for four hours at 37 degrees with shaking. Nalidixic acid (50 ul of a 1 mg/ml aqueous solution) was added and the cultures were incubated at 37 degrees for an additional three hours. Cultures were examined by phase-contrast microscopy for the presence of cytoplasmic inclusion bodies and by SDS polyacrylamide gel electrophoresis for the presence of the expressed protein. Three cultures expressing the most angiostatin were selected for further characterization. Plasmid DNA was isolated from clones C10, D5, and D7 and was analyzed by DNA sequencing. The sequences of the cDNAs in these plasmids are shown in 24643-C10.seq, 24643-D5.seq and 24643-D7.seq (SEQID # 13, SEQID # 14, and SEQID # 15, respectively). Each cDNA encoded the expected protein but contained various silent mutations at the third positions of codons two through eight. These mutations were caused by the redundancies in the 5' PCR primer (angec-s2.seq, SEQID # 2), and the particular mutations in these clones result in high level expression of angiostatin. Each cDNA differed from the others and from the wild-type sequence at one or more of these positions.

Example 3: Construction of pMON24644 and selection of strains producing high levels of angiostatin K2-3

A cDNA encoding angiostatin K2-3 was synthesized by polymerase chain reaction using the oligonucleotide primers anged-s3 (anged-s3.seq, SEQID # 3) and anged-n3 (anged-n3.seq, SEQID # 6) and using the human plasminogen cDNA (pMON24624.seq, SEQID # 9) as template. The resulting cDNA was cleaved with restriction endonuclease *A*/III and *Hind*III and was ligated into the *E. coli* expression vector pMON5723 which had been digested with *Nco*I and *Hind*III. A portion of the ligation reaction was used to transform *E. coli* strain DH10B to spectinomycin resistance

Plasmid DNA was isolated from the DH10B bacterial colonies used to transform *E. coli* strain MON105 to spectinomycin resistance. Individual antibiotic resistant colonies were inoculated into 1 ml of Luria Broth with 75 ug/ml spectinomycin and grown over night at 37 degrees with shaking. Fifty microliters of the overnight culture were inoculated into one ml of M9 medium with casamino acids and grown for four hours at 37 degrees with shaking. Nalidixic acid (50 ul of a 1 mg/ml aqueous solution) was added and the cultures were incubated at 37 degrees for an additional three hours. Cultures were examined by phase-contrast microscopy for the presence of cytoplasmic inclusion bodies and by SDS polyacrylamide gel electrophoresis for the presence of the expressed protein. Three cultures expressing the most angiostatin were selected for further characterization. Plasmid DNA was isolated from clones E7, E9, and F8, and was analyzed by DNA sequencing. The sequences of the cDNAs in these plasmids are shown in 24644-E7.seq, 24644-E9.seq and 24644-F8.seq (SEQID # 16, SEQID # 17, and SEQID # 18, respectively). Each cDNA encoded the expected protein but also contained mutations which altered one or more amino acids. Each cDNA also contained various silent mutations at the third positions of codons two through eight. These mutations were caused by the redundancies in the 5' PCR primer (anged-s3.seq, SEQID # 3), and the particular mutations in these clones result in high level expression of angiostatin. Each cDNA differed from the others and from the wild-type sequence at one or more of these positions.

Example 4: Construction of pMON24645 and selection of strains producing high levels of angiostatin K2-4

A cDNA encoding angiostatin K2-4 was synthesized by polymerase chain reaction using the oligonucleotide primers anged-s3 (anged-s3.seq, SEQID # 3) and

22
angec-n2 (angec-n2.seq, SEQID # 5) and using the human plasminogen cDNA
(pMON24624.seq, SEQID # 9) as template. The resulting cDNA was cleaved with
restriction endonuclease *Afl*III and *Hind*III and was ligated into the *E. coli* expression
vector pMON5723 which had been digested with *Nco*I and *Hind*III. A portion of the
ligation reaction was used to transform *E. coli* strain DH10B to spectinomycin
resistance

Plasmid DNA was isolated from the DH10B bacterial colonies and used to
transform *E. coli* strain MON105 to spectinomycin resistance. Individual antibiotic
resistant colonies were inoculated into 1 ml of Luria Broth with 75 ug/ml
spectinomycin and grown over night at 37 degrees with shaking. Fifty microliters of
the overnight culture were inoculated into one ml of M9 medium with casamino acids
and grown for four hours at 27 degrees with shaking. Nalidixic acid (50 ul of a 1
mg/ml aqueous solution) was added and the cultures were incubated at 37 degrees for
an additional three hours. Cultures were examined by phase-contrast microscopy for
the presence of cytoplasmic inclusion bodies and by SDS polyacrylamide gel
electrophoresis for the presence of the expressed protein. Three cultures expressing
the most angiostatin were selected for further characterization. Plasmid DNA was
isolated from clones G3, H5, and H6, and was analyzed by DNA sequencing. The
sequences of the cDNAs in these plasmids are shown in 24645-G3.seq, 24645-H5.seq
and 24645-H6.seq (SEQID # 19, SEQID # 20, and SEQID # 21, respectively). Each
cDNA encoded the expected protein but clone H5 contained a mutation which altered
one amino acid. Each cDNA also contained various silent mutations at the third
positions of codons two through eight. These mutations were caused by the
redundancies in the 5' PCR primer (angec-s3.seq, SEQID # 3), and the particular
mutations in these clones result in high level expression of angiostatin. Each cDNA
differed from the others and from the wild-type sequence at one or more of these
positions.

Example 5: Construction of pMON24646 and selection of strains producing high levels of angiostatin K1-3

A cDNA encoding angiostatin K1-3 was synthesized by polymerase chain
reaction using the oligonucleotide primers angec-s2 (angec-s2.seq, SEQID # 2) and
angec-r3 (angec-n3.seq, SEQID # 6) and using the human plasminogen cDNA
(pMON24624.seq, SEQID # 9) as template. The resulting cDNA was cleaved with
restriction endonuclease *Afl*III and *Hind*III and was ligated into the *E. coli* expression
vector pMON5723 which had been digested with *Nco*I and *Hind*III. A portion of the

ligation reaction was used to transform *E. coli* strain DH10B to spectinomycin resistance

Plasmid DNA was isolated from the DH10B bacterial colonies and used to transform *E. coli* strain MON105 to spectinomycin resistance. Individual antibiotic resistant colonies were inoculated into 1 ml of Luria Broth with 75 ug/ml spectinomycin and grown over night at 37 degrees with shaking. Fifty microliters of the overnight culture were inoculated into one ml of M9 medium with casamino acids and grown for four hours at 37 degrees with shaking. Nalidixic acid (50 ul of a 1 mg/ml aqueous solution) was added and the cultures were incubated at 37 degrees for an additional three hours. Cultures were examined by phase-contrast microscopy for the presence of cytoplasmic inclusion bodies and by SDS polyacrylamide gel electrophoresis for the presence of the expressed protein. Six cultures expressing the most angiostatin were selected for further characterization. Plasmid DNA was isolated from clones A7, A9, A10, A12, B1, and B10, and was analyzed by DNA sequencing. The sequences of the cDNAs in these plasmids are shown in 24646-A7.seq, 24646-A9.seq and 24646-A10, 24646-A12, 24646-B1, and 24646-B10 (SEQID # 22, SEQID # 23, SEQID # 24, SEQID # 25, SEQID # 26, and SEQID # 27, respectively). Each cDNA encoded the expected protein. Each cDNA also contained various silent mutations at the third positions of codons two through eight. These mutations were caused by the redundancies in the 5' PCR primer (angec-s2.seq, SEQID # 2), and the particular mutations in these clones result in high level expression of angiostatin. Each cDNA differed from the others and from the wild-type sequence at one or more of these positions.

Example 6: Construction of pMON24649 and of strains producing high levels of angiostatin K1

A cDNA encoding angiostatin K1 (24649.seq, SEQID # 28) was synthesized by polymerase chain reaction using the oligonucleotide primers angec-sk1 (angec-sk1.seq, SEQID # 7) and angec-n4 (angec-n4.seq, SEQID # 8) and using the cDNA in pMON24646-A7 (SEQID # 22) as template. The 5' PCR primer (angec-sk1) contained the same silent mutations in the first eight codons that resulted in high level expression of angiostatin K1-3 from pMON24646-A7. The resulting cDNA was cleaved with restriction endonucleases *NcoI* and *HindIII* and was ligated into the *E. coli* expression vector pMON5723 which had also been digested with *NcoI* and *HindIII*. A portion of the ligation reaction was used to transform *E. coli* strain DH10B to spectinomycin resistance. Plasmid DNA was isolated from the bacterial colonies and analyzed by restriction endonuclease cleavage and DNA sequencing.

Plasmid DNA was then used to transform *E. coli* strain MON105 to spectinomycin resistance. Individual colonies were grown over night in Luria Broth containing 75 ug/ml spectinomycin. One ml of the over night culture was used to inoculate 20 ml of M9 medium containing casamino acids. After about three hours of growth at 37 degrees, Nalidixic acid was added and the cultures incubated for a further three hours at 37 degrees. Cultures were screened by phase contrast microscopy for the presence of cytoplasmic inclusion bodies and by SDS polyacrylamide gel electrophoresis for expression of the protein. Unlike other forms of angiostatin, K1 protein was not incorporated into inclusion bodies, but rather remained soluble in the cytoplasm of the bacteria. Nevertheless, the mutations in the first eight codons in pMON24649 conferred high level expression of K1 as they had for K1-3 expression from pMON13646-A7.

Example 7: Method for refolding angiostatin

All steps are at 4°C. Angiostatin K1-4 inclusion bodies were dissolved in 6M urea, 5 mM DTT, 50 mM Bis-Tris Propane, pH 10.8 at 1.2 mg/ml. This solution was stirred for 2 hours. A stock solution of 1M 6-amino hexanoic acid was then added to achieve a 6 mM concentration, followed by addition of cystine (stock 0.2M, pH 10.5) to 6 mM. This was stirred for 5 minutes. The K1-4 angiostatin was then diluted to 0.2 mg/ml by addition of 75 mM Bis-Tris propane, pH 10.5. It was then stirred for 72 hours to complete the refolding of the protein. Refolded angiostatin was purified by lysine-Sepharose chromatography and Q-Sepharose-HP chromatography. This purified angiostatin was assayed for inhibition of cell migration as shown in figures 3-5.

All references, patents, or applications cited herein are incorporated by reference in their entirety, as if written herein.

Tables

Table 1: Strains

Designation	Description or Genotype	Reference/Source
DH10B	F ⁻ <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) ϕ dlacZ Δ M15 Δ lacX74 <i>deoR recA1</i> <i>endA1 araD139</i> Δ (<i>ara leu</i>)7697 <i>galU galK</i> λ <i>rpsL nupG</i>	Life Technologies, Rockville, Maryland
MON105 (ATCC # 55204)	F ⁺ , λ IN (<i>rrnD</i> , <i>rrnE</i>)1, <i>rpoD</i> ⁺ , <i>rpoH</i> 358	Obukowicz et al., <i>Appl. and Envir.</i> <i>Micr.</i> , 58: 1511-1523, 1992
JM101 (ATCC # 33876)	<i>delta</i> (<i>pro lac</i>), <i>supE</i> , <i>thi</i> , F'(<i>traD</i> 36, <i>proA</i> ⁺ <i>B</i> ⁺ , <i>lacI</i> ^q , <i>lacZ</i> Δ M15)	Yanisch-Perron et al., <i>Gene</i> , 33: 103-119, 1985

Table 2: Plasmids

Plasmid	SEQ ID NO.	Marker	Description	Source
pMON5723		Spec	Generic Spec-resistant <i>E. coli</i> expression vector containing the <i>E. coli recA1</i> promoter and G10L ribosome binding site	(Olins and Rangwala, <i>Methods Enzymol.</i> 185 (Gene Expression Technol.): 115-119, 1990)
pMON6875		Spec	Generic Spec-resistant <i>E. coli</i> expression vector containing the <i>E. coli recA1</i> promoter and G10L ribosome binding site	(Olins and Rangwala, <i>Methods Enzymol.</i> 185 (Gene Expression Technol.): 115-119, 1990)
pMON24624	#9	Amp	Human plasminogen, including the native plasminogen signal sequence, N-terminal peptide and kringle 1-5. <i>Bam</i> HI fragment in pMON3360B. (Amino acids -19-537 of plasminogen, nucleotides 55-1752)	This work
pMON24642-A12, -B5, -B6	#10, #11, #12	spec	Human angiotensinogen(E). <i>Afl</i> III/ <i>Hind</i> III fragment in pMON5723. The 5' end has been mutagenized for increased expression. (Amino acids 74-470 of plasminogen, nucleotides 331-1464)	This work

pMON24643- C10, D5,-D7	#13, #14, #15	spec	Human angiotatin(S) (K1-4) <i>AflIII/HindIII</i> fragment in pMON5723. (Amino acids 79-440 of plasminogen, nucleotides 346-1431)	This work
pMON24644- E7, -E9, -F8	##16, #17, #18	spec	K2-3 fragment of human angiotatin(S). <i>AflIII/HindIII</i> fragment in pMON5723. (Amino acids 163-344 of plasminogen, nucleotides 598-1143)	This work
pMON24645- G3, -H5, -H6	#19, #20, #21	spec	K2-4 fragment of human angiotatin(S). <i>AflIII/HindIII</i> fragment in pMON5723. (Amino acids 163-440 of plasminogen, nucleotides 598-1431)	This work
pMON24646- A7, -A9, -A10, -A12, -B1, - B10	##22, #23, #24, #25, #26, #27	spec	K1-3 fragment of human angiotatin(S). <i>AflIII/HindIII</i> fragment in pMON5723. (Amino acids 79-344 of plasminogen, nucleotides 346-1143)	This work
pMON24649	#28	spec	K1 fragment of human angiotatin(S). <i>NcoI/HindIII</i> fragment in pMON5723. (Amino acids 79-184 of plasminogen, nucleotides 1-269 of pMON24646- A7.seq)	This work
pMON24650	#29	Amp	K1 fragment of human angiotatin(S). <i>NcoI/HindIII</i> fragment in pMON3633. (Amino acids 79-184 of plasminogen)	This work

Table 3: SEQ ID Correlation Table

SEQ ID NO:	Sequence	Description	Name
1	gatcacatgtchctbtttgaraaraa rgtbtatctctcagagtgcgaag	oligo K1E	anged-s1.seq
2	gatcacatgtchgtbtatctbtchga rtgyaagactgggaatggaaag	oligo K1	angeds2.seq
3	gatcacatgtchgargargartgyat gcaytgcagtgagaaaaactat	oligo K2	angeds3.seq
4	gatcaagctttttattaaagcaggaca acaggcgg	oligo K4E	anged-n1.seq
5	gatcaagctttttattacgttctgtt cctgagca	oligo K4	anged-n2.seq
6	gatcaagctttttattaagccaattgt tccgtgga	oligo K3	anged-n3.seq
7	catgccatggcagtttatctttcaga gtgt	oligo K1	hanged-sk1.seq
8	gatcaagctttttattattcctcttca cactcaag	oligo K1	anged-n4.seq
9		human plasminogen, hPlgn	24624.seq
10		K1-4E/5723	24642-a12.seq
11		K1-4E/5723	24642-b5.seq
12		K1-4E/5723	24642-b6.seq
13		K1-4/5723	24643-C10.seq
14		K1-4/5723	24643-d5.seq
15		K1-4/5723	24643-d7.seq
16		K2-3/5723	24644-e7.seq
17		K2-3/5723	24644-e9.seq
18		K2-3/5723	24644-f8.seq
19		K2-4/5723	24645-g3.seq
20		K2-4/5723	24645-h5.seq
21		K2-4/5723	24645-h6.seq
22		K1-3/5723	24646-a7.seq
23		K1-3/5723	24646-a9.seq

24	K1-3/5723	24646-a19.seq
25	K1-3/5723	24646-a12.seq
26	K1-3/5723	24646-b1.seq
27	K1-3/5723	24646-b10.seq
28	K1/5723 (24646-A7 5'end)	24649.seq
29	K1/3633 (same cDNA as 24649)	24650.seq

What is claimed is.

1. A method of producing angiostatin comprising the steps of:
 - (a) culturing host cells that are expressing an angiostatin gene;
 - 5 (b) recovering the product of said gene expression;
 - (c) solubilizing said gene product at an elevated pH;
 - (d) refolding said solubilized gene product at an elevated pH; and
 - (e) isolating properly folded forms of said gene product.
2. A method of producing angiostatin comprising the steps of:
 - 10 (a) culturing host cells that are expressing an angiostatin gene;
 - (b) recovering the product of said gene expression;
 - (c) refolding said gene product at an elevated pH; and
 - (d) isolating properly folded forms of said gene product.
3. The method of claim 1 or claim 2 wherein said elevated pH is from
15 about pH 9 to about pH 11.5.
4. The method of claim 3 wherein said elevated pH is from about pH 10 to about pH 11.
5. The method of claim 4 wherein said elevated pH is about pH 10.5.
6. The method of claim 1 or claim 2 wherein said gene product is present
20 at a concentration of from about 1 to about 20 mg/ml during said solubilization step.
7. The method of claim 6 wherein said gene product is present at a concentration of about 6 mg/ml during said solubilization step.
8. The method of claim 1 or claim 2 wherein said gene product is present
25 at a concentration of from about 0.1 to about 5 mg/ml during said

refolding step.

9. The method of claim 8 wherein said gene product is present at a concentration of about 1 mg/ml during said refolding step.
- 5 10. The method of claim 1 or claim 2 wherein said angiostatin gene product is encoded by a sequence is selected from the group consisting of SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 15; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 18; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 21; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 24; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; SEQ ID NO: 29.
- 10 11. The method of claim 1 or claim 2 wherein urea is present at a concentration of from about 4 M to about 10 M during said solubilization step.
- 15 12. The method of claim 11 wherein urea is present at a concentration of about 6 M during said solubilization step.
13. The method of claim 1 or claim 2 wherein urea is present at a concentration of from about 0.5 M to about 3 M during said refolding step.
- 20 14. The method of claim 13 wherein urea is present at a concentration of about 1 M during said refolding step.
15. The method of claim 1 or claim 2 wherein guanidine hydrochloride is present at a concentration of from about 3 M to about 8 M during said solubilization step.
- 25 16. The method of claim 15 wherein guanidine hydrochloride is present at a concentration of about 5 M during said solubilization step.
17. The method of claim 1 or claim 2 wherein guanidine hydrochloride is present at a concentration of from about 0.2 M to about 2 M during said refolding step.

18. The method of claim 17 wherein guanidine hydrochloride is present at a concentration of about 0.5 M during said refolding step.
19. A method of claim 1 or 2 carried out in a presence of a reducing agent capable of reducing disulfide linkages to sulfhydryl groups.
- 5 20. A method of claim 19 wherein said reducing agent is selected from the group consisting of DTT, BME, cysteine, and reduced glutathione.
21. The method of claim 20 wherein DTT is present at a concentration of from about 2 mM to about 10 mM during said solubilization step.
- 10 22. The method of claim 21 wherein DTT is present at a concentration of about 6 mM during said solubilization step.
23. The method of claim 20 wherein DTT is present at a concentration of from about 0.5 mM to about 2 mM during said refolding step.
24. The method of claim 23 wherein DTT is present at a concentration of about 1 mM during said refolding step.
- 15 25. The method of claim 20 wherein reduced glutathione is present at a concentration of from about 5 mM to about 20 mM during said solubilization step.
26. The method of claim 25 wherein reduced glutathione is present at a concentration of about 10 mM during said solubilization step.

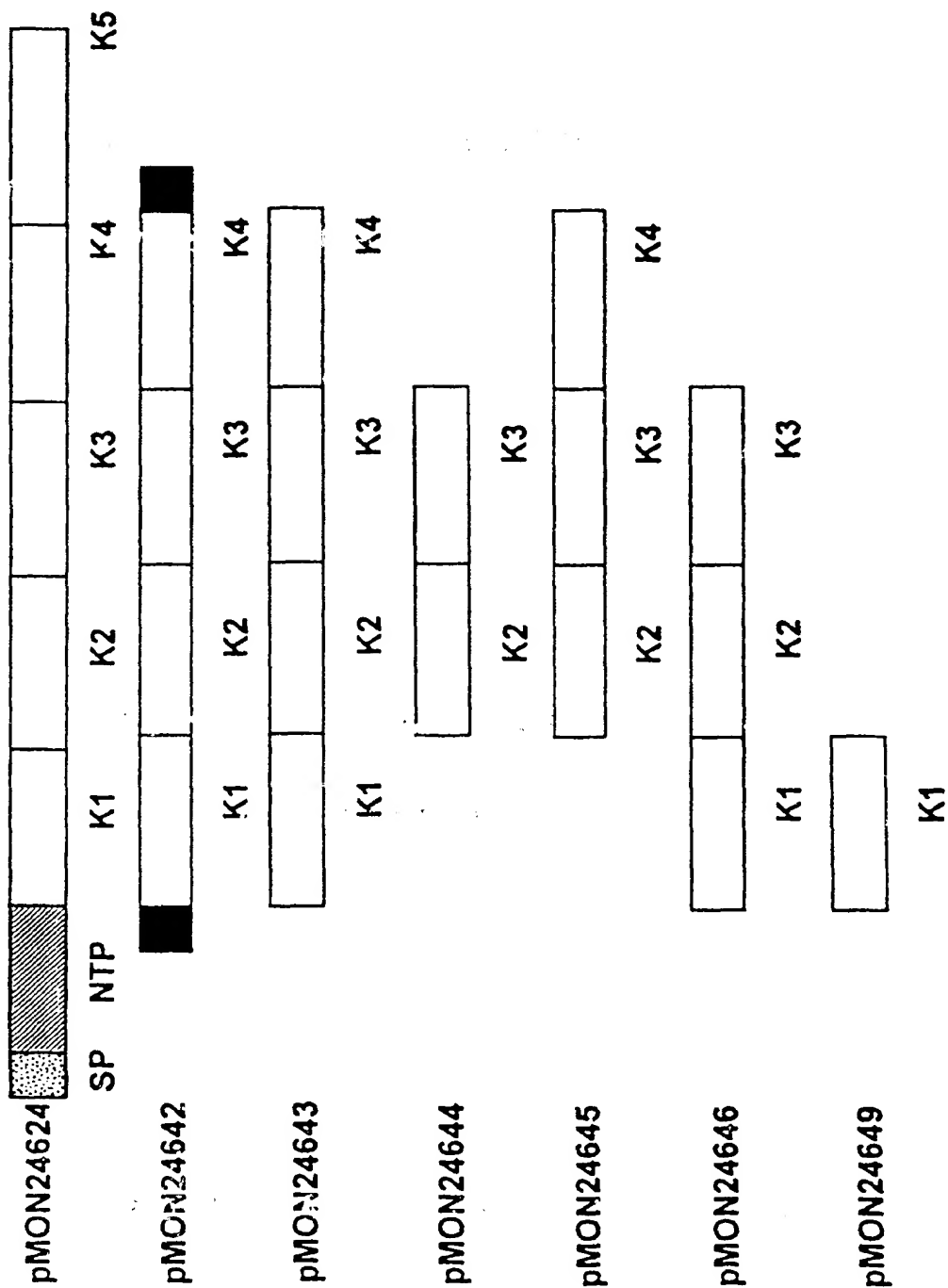
27. The method of claim 20 wherein reduced glutathione is present at a concentration of from about 1 mM to about 4 mM during said refolding step.
28. The method of claim 27 wherein reduced glutathione is present at a concentration of about 2 mM during said refolding step.
29. The method of claim 20 wherein cysteine is present at a concentration of from about 5 mM to about 20 mM during said solubilization step.
30. The method of claim 29 wherein cysteine is present at a concentration of about 10 mM during said solubilization step.
31. The method of claim 20 wherein cysteine is present at a concentration of from about 1 mM to about 4 mM during said refolding step.
32. The method of claim 31 wherein cysteine is present at a concentration of about 2 mM during said refolding step.
33. The method of claim 1 or claim 2 wherein an agent capable of enhancing the interchange of disulfide bonds is present during said refolding step.
34. The method of claim 33 wherein said agent is selected from cystine and oxidized glutathione.
35. The method of claim 34 wherein cystine is present at a concentration of from about 0.2 mM to about 5 mM during said refolding step.
36. The method of claim 35 wherein cystine is present at a concentration of about 1 mM during said refolding step.
37. The method of claim 1 or 2 wherein said refolding step is carried out in the presence of a ligand capable of binding kringle domains of said angiostatin gene product.
38. The method of claim 37 wherein said ligand capable of binding kringle domains of said angiostatin gene product is selected from the group consisting of 6-amino hexanoic acid, 7-aminoheptanoic acid, 5-amino pentanoic acid, and 4-amino butyric acid.

39. The method of claim 38 wherein 6-amino hexanoic acid is present at a concentration of from about 0.2 mM to about 20 mM during said refolding step.
40. The method of claim 39 wherein 6-amino hexanoic acid is present at a concentration of about 1 mM during said refolding step.
41. The method of claim 1 or claim 2 wherein disulfide bonds are formed through air oxidation during said refolding step.
42. The method of claim 41 wherein said air oxidation step is carried out from about 12 hours to about 96 hours.
43. The method of claim 42 wherein said air oxidation step is carried out from about 24 hours to about 72 hours.
44. The method of claim 43 wherein said air oxidation step is carried out for about 60 hours.
45. The method of claim 1 or claim 2 which further includes the step of purifying the angiostatin by a process selected from the group consisting of Lysine-sepharose chromatography, ion-exchange chromatography, hydrophobic interaction chromatography and RP-HPLC.
46. The method of claim 1 or claim 2 wherein said angiostatin gene is comprised of DNA coding for animal angiostatin.
47. The method of claim 1 or claim 2 wherein said angiostatin gene is comprised of DNA coding for human angiostatin.
48. The method of claim 1 or claim 2 wherein said angiostatin gene is selected from the group consisting of SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 15; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 18; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 21; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 24; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; SEQ ID NO: 29.

5

49. A polypeptide encoded by a sequence selected from the group consisting of SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 15; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 18; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 21; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 24; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; SEQ ID NO: 29.

FIG. 1



2/6

FIG. 2

SOLUBILIZATION OF IBS IN

6M UREA, 5mM DTT, pH 10.5

MIX 2 HOURS



ADDITION OF 6 mM CYSTINE, 6 mM 6-AHA

MIX 5 MINUTES

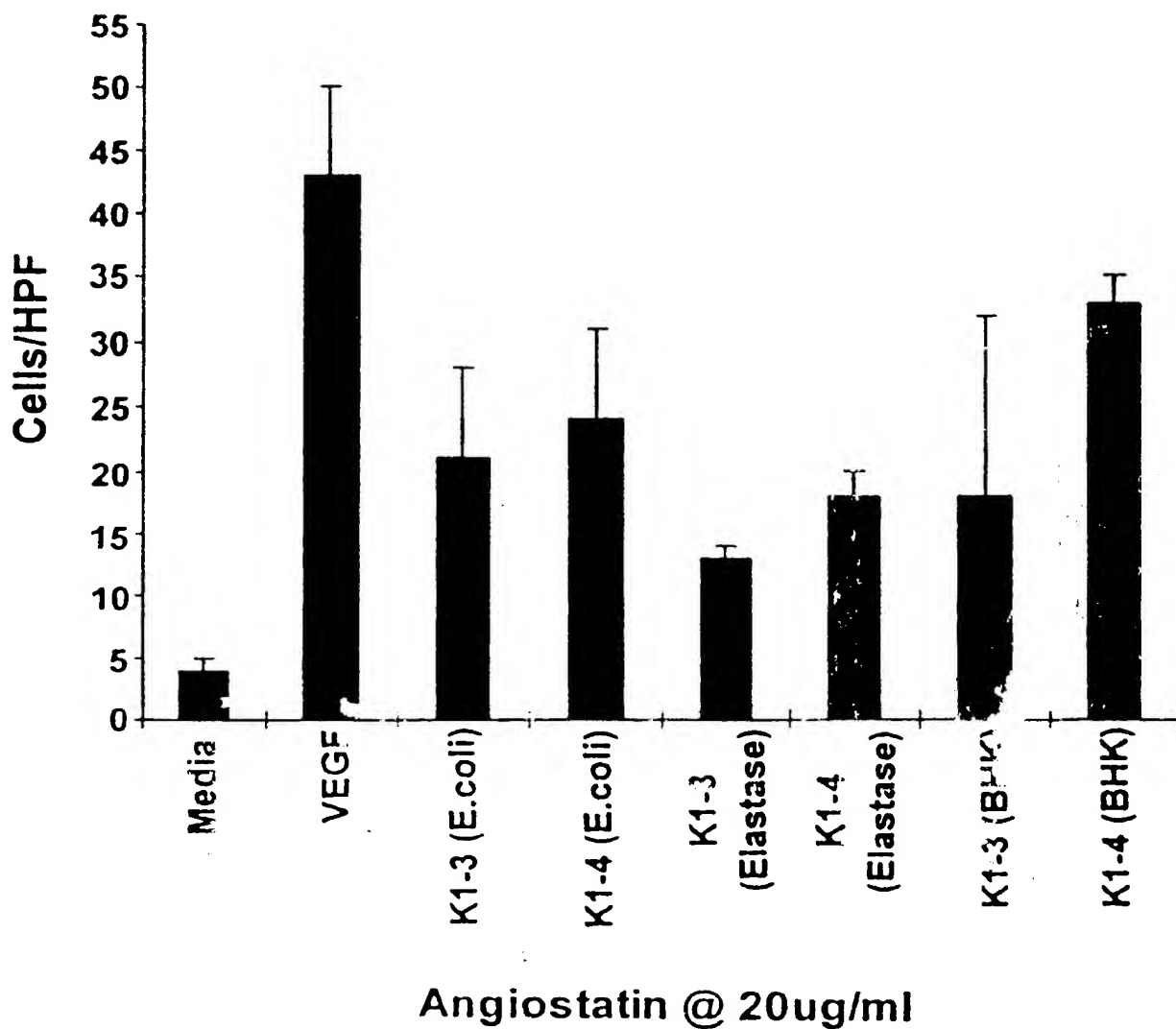


DILUTE 6-FOLD AT pH 10.5

STIR 72 HOURS

FIG. 3

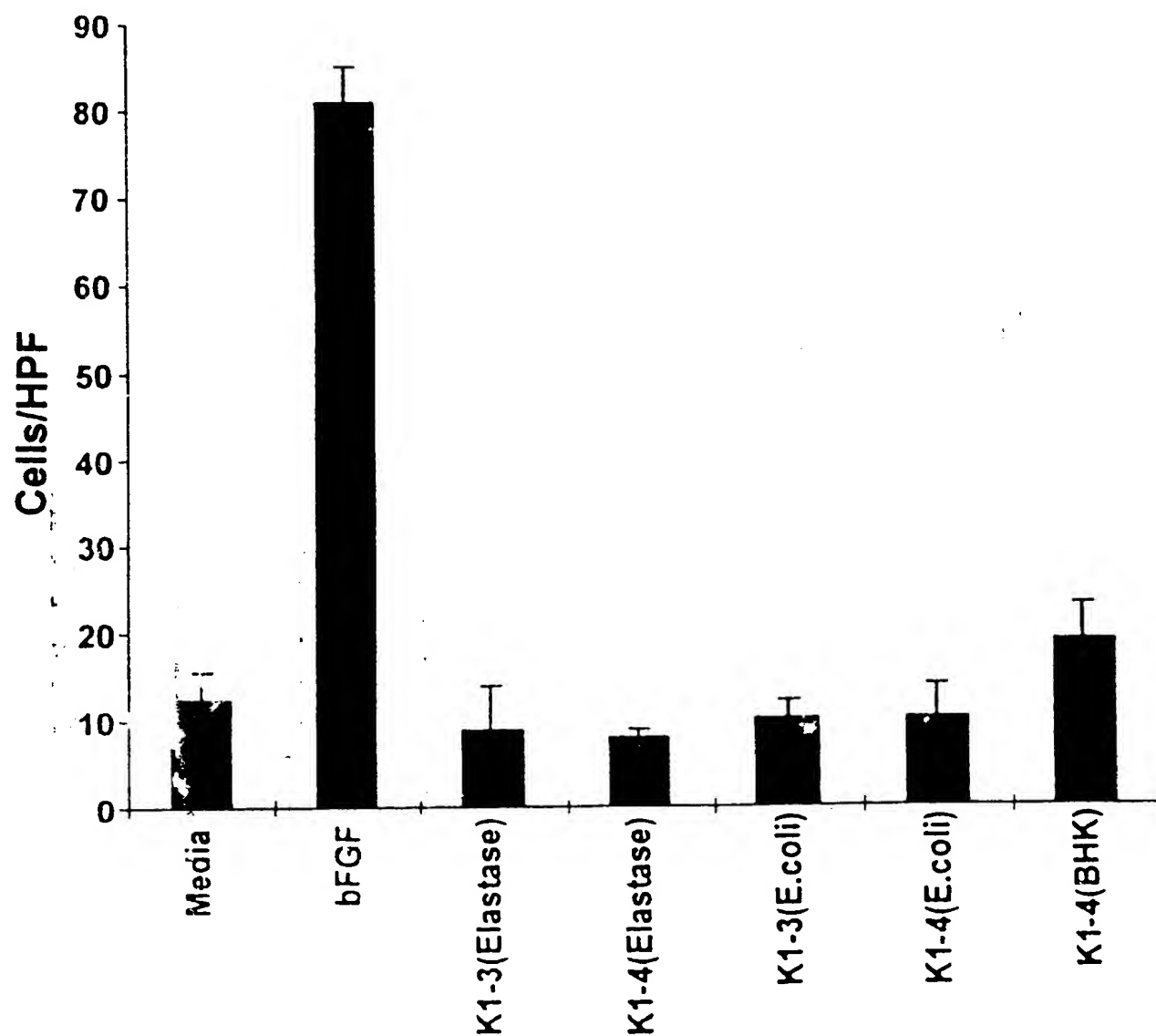
Inhibition of HMEC Migration



4/6

FIG. 4

Inhibition of BAEC Migration

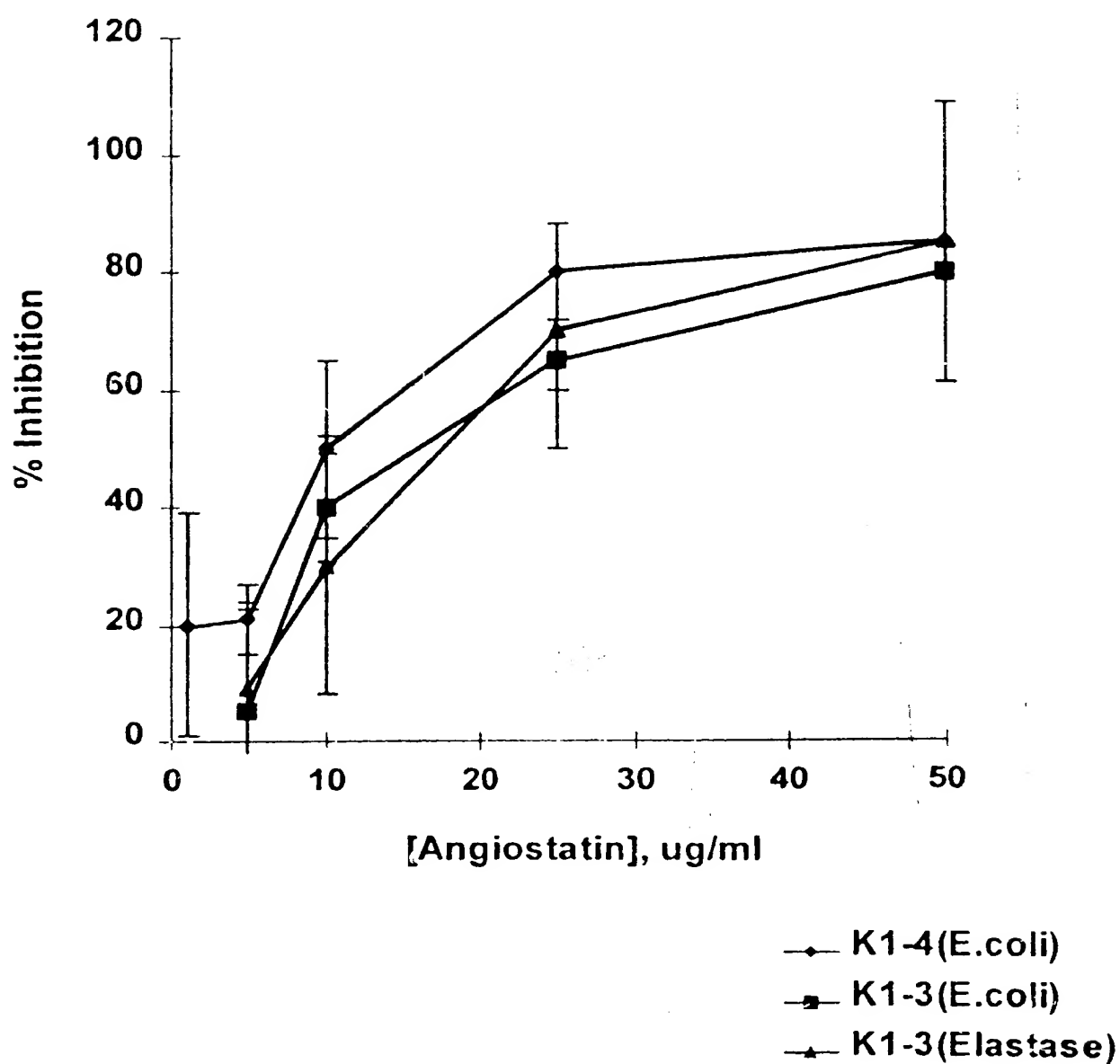


Angiostatin @ 20ug/ml

SUBSTITUTE SHEET (RULE 26)

FIG. 5

Inhibition of HMEC Migration



6/6

FIG. 6A

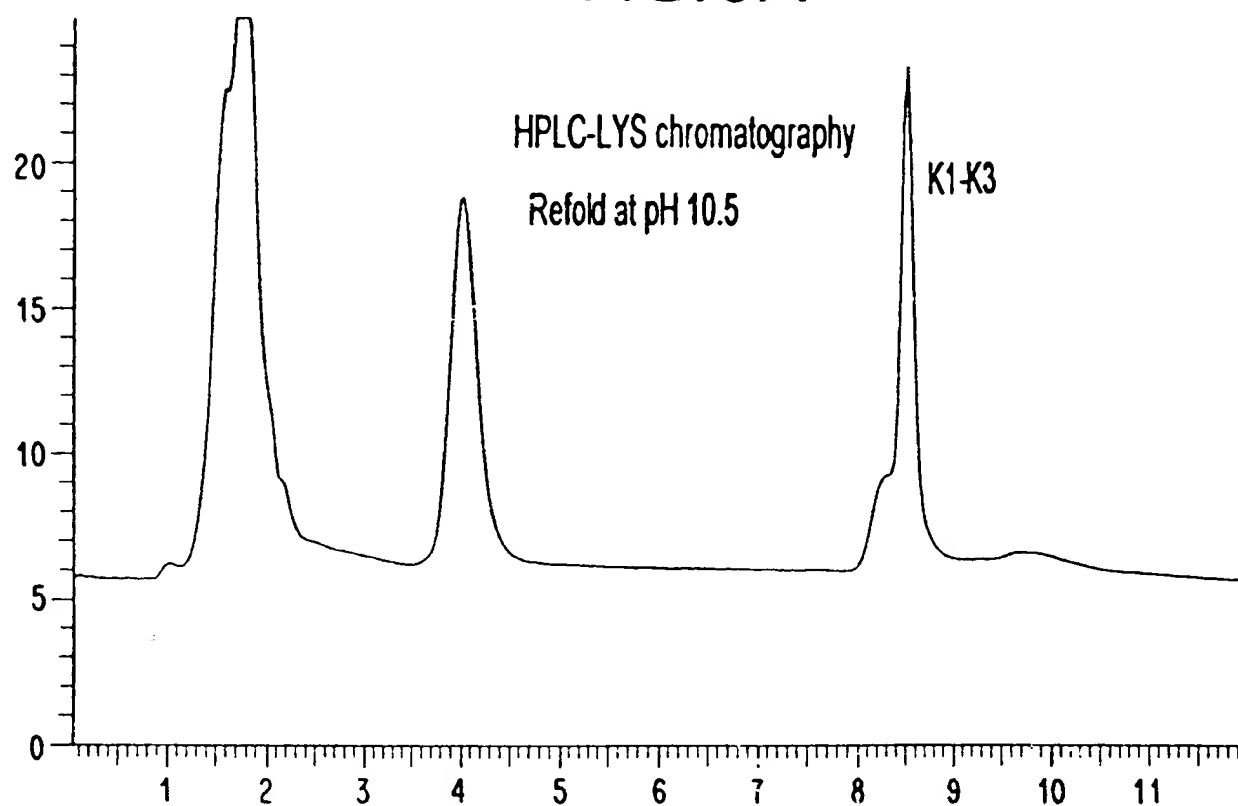
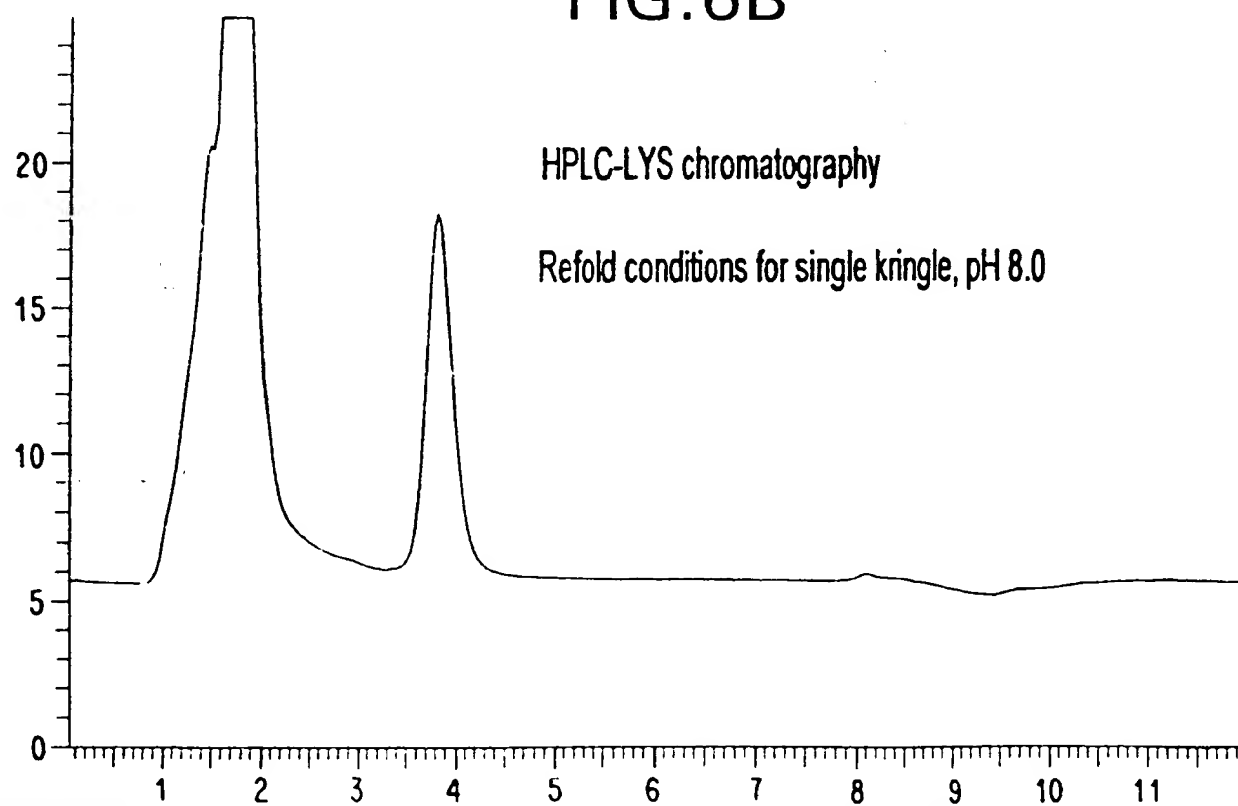


FIG. 6B



Sequence Listing

SEQUENCE LISTING

5 <110> Violand, B.N.
 Polazzi, J.O.
 Casperson, G.F.

 <120> Method of refolding of angiostatin

10

 <130> C_3044-0

 <150> 60/071,247

 <151> 1998-01-12

15

 <160> 29

 <170> FastSEQ for Windows Version 3.0

20

 <210> 1

 <211> 48

 <212> DNA

 <213> human

25

 <400> 1

2

gatcacatgt chctbtttga raaraargtb tatctctcag agtgcaag

48

<210> 2

<211> 48

5

<212> DNA

<213> human

<400> 2

gatcacatgt chgtbtatct btchgartgy aagactggga atggaaag

48

10

<210> 3

<211> 48

<212> DNA

<213> human

15

<400> 3

gatcacatgt chgargarga rtgyatgcay tgcagtggag aaaactat

48

<210> 4

20

<211> 34

<212> DNA

<213> human

<400> 4

25

gatcaagctt ttattaaagc aggacaacag gcgg

34

<210> 5

<211> 34

<212> DNA

<213> human

5

<400> 5

gatcaagctt ttattacgct tctgttctcg agca

34

<210> 6

10

<211> 34

<212> DNA

<213> human

<400> 6

15

gatcaagctt ttattaagcc aattgttccg tgga

34

<210> 7

<211> 30

<212> DNA

20

<213> human

<400> 7

catgccatgg cagtttatct ttcagagtgt

30

25

<210> 8

<211> 34

4

<212> DNA

<213> human

<400> 8

5 gatcaagctt ttattattcc ttttcacact caag

34

<210> 9

<211> 1719

<212> DNA

10 <213> human

<400> 9

	ggatccgccca ccatggaaca taaggaagtg gttcttctac ttcttttatt tctgaaatca	60
	ggtaaggag agcctctgga tgactatgtg aataccagg gggcttcaact gtccagtgtc	120
15	actaagaagc agctgggagc aggaagtata gaagaatgtg cagcaaatg tgaggaggac	180
	gaagaattca cctgcagggc attccaatat cacagtaaag agcaacaatg tgtgataatg	240
	gctgaaaaca ggaagtcttc cataatcatt aggatgagag atgtagtgtt atttgaaaag	300
	aaagtgtatc tctcagagtg caagactggg aatggaaaga actacagagg gacgatgtcc	360
	aaaacaaaaa atggcatcac ctgtcaaaaa tggagtcca cttctcccca cagacctaga	420
20	ttctcacctg ctacacaccc ctccagaggga ctggaggaga actactgcag gaatccagac	480
	aacgatccgc aggggccctg gtgctatact actgatccag aaaagagata tgactactgc	540
	gacattcttg agtgtgaaga ggaatgtatg cattgcagtg gagaaaacta tgacggcaaa	600
	atttccaaga ccatgtcttg actggaatgc caggcctggg actctcagag cccacacgct	660
	catggataca ttcttccaa atttccaaac aaqaacctga agaagaatta ctgtcgtaac	720
25	cccgataggg agctgcggcc ttggtgtttc accaccgacc ccaacaagcg ctgggaactt	780
	tgtgacatcc cccgctgcac aacacctcca ccatcttctg gtccaccta ccagtgtctg	840

aagggaacag gtgaaaacta tgcggggaat gtggctgtta ccgtgtccgg gcacacctgt 900
 cagcactgga gtgcacagac cctcacaca cataacagga caccagaaaa cttccctgc 960
 aaaaatttgg atgaaaacta ctgccgaat cctgacggaa aaagggcccc atggtgccat 1020
 acaaccaaca gccaaagtgcg gtgggagtag tgtaagatac cgtcctgtga ctctcccca 1080
 5 qtatccacgg aacaattggc tcccacagca ccacctgagc taacctctgt ggtccaggac 1140
 tgctaccatg gtgatggaca gagctaccga ggcacatcct ccaccaccac cacaggaaaag 1200
 aagtgtcagt cttggtcatc tatgacacca caccggcacc agaagacccc agaaaactac 1260
 ccaaattgtg gcctgacaat gaactactgc aggaatccag atgccgataa aggcccttg 1320
 tgttttacca cagaccccag cgtcaggtgg gagtactgca acctgaaaaa atgctcagga 1380
 10 acagaagcga gtgtttagc acctccgct gtgtcctgc ttccagatgt agagactcct 1440
 tccgaagaag actgtatgtt tgggaatggg aaaggatacc gaggcaagag ggcgaccact 1500
 gtactggga cgccatgccg ggactgggct gcccaggagc cccatagaca cagcattttc 1560
 actccagaga caaatccacg ggcgggtctg gaaaaaatt actgccgtac cctgatggt 1620
 gatgtaggtg gtccctggtg ctacacgaca aatccaagaa aactttacga ctactgtgat 1680
 15 gtcccwcagt gtgcggcccc ttcatttgat taaggatcc 1719

<210> 10

<211> 1154

<212> DNA

20

<213> human

<400> 10

ccatgtcact ttttgagaag aaggctctatc tctcagagtg caagactggg aatggaaaaga 60
 attacagagg gacgatgtcc aaaacaaaaa atggcatcac ctgtcaaaaa tggagttcca 120
 25 cttctcccca cagacctaga ttctcacctg ctacacaccc ctgagaggga ctggaggaga 180
 actactgcag gaatccagac aacgatccgc aggggccctg gtgctatact actgatccag 240

6

aaaagagata tgactactgc gacattcttg agtgtgaaga ggaatgtatg cattgcagtg 300
 gagaaaaacta tgacggcaaa atttccaaga ccatgtcttg actggaatgc caggcctggg 360
 actctcagag cccacacgct catggataca ttccctccaa atttccaaac aagaacctga 420
 agaagaatta ctgtcgtaac cccgataggg agctgcggcc ttggtgttcc accaccgacc 480
 5 ccaacaagcg ctgggaactt tgtgacatcc cccgtgcac aacacctcca ccatcttctg 540
 gtcccaccta ccagtgtctg aagggaacag gtgaaaacta tcgcgggaat gtggctgtta 600
 ccgtgtccgg gcacacctgt cagcactgga gtgcgcagac cctcacaca cataacagga 660
 caccagaaaa ctccccctgc aaaaatttgg atgaaaacta ctgccgcaat cctgacggaa 720
 aaagggcccc atggtgccat acaaccaaca gccagtgcg gtgggagtac tgtaagatac 780
 10 cgtcctgtga ctctcccca gtatccacgg aacagttggc tcccacagca ccacctgagc 840
 taacccctgt ggtccaggac tgcctaccatg gtgatggaca gagctaccga ggcacatcct 900
 ccaccaccac cacaggaaag aagtgtcagt ctgggtcacc tatgacacca caccggcacc 960
 agaagacccc agaaaactac ccaaagtctg gcctgacaat gaactactgc aggaatccag 1020
 atgccgataa aggccctctg tgttttacc aagaccccag cgtcagggtg gagtactgca 1080
 15 acctgaaaaa atgctcagga acagaagcga gtgttctagc acctccgct gttgtcctgc 1140
 tttaataaaa gctt 1154

<210> 11

<211> 1154

20

<212> DNA

<213> human

<400> 11

25

ccatgtcact gtttgaaaaa aaagtttacc tctcagagtg caaactggg aatggaaga 60
 attacagagg gacgatgtcc aaaacaaaaa atggcatcac ctgtcaaaaa tggagtcca 120
 ctctcccca cagacctaga ttctcacctg ctacacccc ctcagaggga ctggaggaga 180

7

actactgcag gaatccagac aacgatccgc aggggccctg gtgctatact actgatccag 240
 aaaagagata tgactactgc gacattcttg agtgtgaaga ggaatgtatg cattgcagtg 300
 gagaaaacta tgacggcaaa atttccaaga ccatgtcttg actggaatgc caggcctggg 360
 actctcagag cccacacgct catggataca ttcttccaa atttccaaac aagaacctga 420
 5 agaagaatta ctgtcgtaac cccgataggg agctgcggcc ttggtgttcc accaccgacc 480
 ccaacaagcg ctgggaactt tgtgacatcc cccgtgcac aacacctcca ccatctcttg 540
 gtccaccta ccagtgtctg aagggaacag gtgaaaacta tcgcgggaat gtggctgtta 600
 ccgtgtccgg gcacacctgt cagcactgga gtgcacagac cctcacaca cataacagga 660
 caccagaaaa ctccccctgc aaaaatttgg atgaaaacta ctgcgcgaat cctgacggaa 720
 10 aaagggcccc atggtgccat acaaccaaca gccaaagtgcg gtgggagtac tgtaagatac 780
 cgtcctgtga ctctcccca gtatccacgg aacagttggc tccacagca ccacctgagc 840
 taacctctgt ggtccaggac tgctaccatg gtgatggaca gagctaccga ggcacatcct 900
 ccaccaccac cacaggaaag aagtgtcagt cttggtcacc tatgacacca caccggcacc 960
 agaagacccc agaaaactac ccaaagtctg gcctgacaat gaactactgc aggaatccag 1020
 15 atgccgataa agggcccttg tgttttacca cagaccccag cgtcaggtgg gactactgca 1080
 acctgaaaaa atgctcagga acagaagcga gtgttgtagc acctccgctt gttgtcctgc 1140
 ttttaataaaa gctt 1154

<210> 12

20

<211> 1154

<212> DNA

<213> human

<400> 12

25

ccatgtccct ttttgaaaaa aaagtttatc tctcagagtg caagactggg aatggaaaga 60
 attacagagg gacgatgtcc aaaacaaaaa atggcatcac ctgtcaaaaa tggagtcca 120

	cttctcccca cagacctaga ttctcactg ctacacaccc ctcagagggg ctggaggaga	180
	actactgcag gaatccagac aacgatccgc aggggccctg gtgctatact actgatccag	240
	aaaagagata tgactactgc gacattcttg agtgtgaaga ggaatgtatg cattgcagtg	300
	gagaaaacta tgacggcaaa atttccaaga ccatgtctgg actggaatgc caggcctggg	360
5	actctcagag cccacacgct catggataca ttcttccaa atttccaaac aagaacctga	420
	agaagaatta ctgtcgtaac cccgataggg agctgcggcc ttggtgtttc accaccgacc	480
	ccaacaagcg ctgggaactt tgtgacatcc cccgtgcac aacacctcca ccatcttctg	540
	gtcccaacta ccagtgtctg aagggaacag gtgaaaacta tcgcgggaat gtggctgtta	600
	ccgtgtccgg gcacacctgt cagcactgga gtgcacagac cctcacaca cataacagga	660
10	caccagaaaa ctccccctgc aaaaatttgg atgaaaacta ctgccgaat cctgacggaa	720
	aaagggcccc atggtgccat acaaccaaca gccaaagtgcg gtgggagtac tgtaagatac	780
	cgctctgtga ctctcccca gtatccacgg aacagttggc tcccacagca ccacctgagc	840
	taaccctgtg ggtccaggac tgctaccatg gtgatggaca gagctaccga ggcacatcct	900
	ccaccaccac cacaggaaag aagtgtcagt ctgggtcatc tatgacacca caccggcacc	960
15	agaagacccc agaaaactac ccaaattgctg gcctgacaat gaactactgc aggaatccag	1020
	atgccgataa agggccctgg tgttttacca cagaccccag cgtcaggtgg gagtactgca	1080
	acctgaaaaa atgctcagga acagaagcga gtgttgtagc acctccgcct gttgtcctgc	1140
	tttaataaaa gctt	1154
20	<210> 13	
	<211> 1106	
	<212> DNA	
	<213> human	
25	<400> 13	
	ccatgtcagt gtatctgtcc gagtgtaaga ctgggaatgg aaagaattac agagggacga	60

5
10
15

tgtccaaaac aaaaaatggc atcacctgtc aaaaaatggag ttccatttct cccacagac 120
ctagatttct acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc 180
cagacaacga tccgcagggg ccttggtgct atactactga tccagaaaag agatatgact 240
actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg 300
gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagccac 360
acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc 420
gtaaccccgga tagggagctg cggccttggg gtttcaccac cgacccaac aagcgctggg 480
aactttgtga catccccgc tgcacaacac ctccaccatc ttctgggtcc acctaccagt 540
gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca 600
cctgtcagca ctggagtga cagaccctc acacacataa caggacacca gaaaacttcc 660
cctgcaaaaa ttggatgaa aactactgcc gcaatctga cggaaaaagg gcccctggt 720
gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct 780
ccccagtatc caggaacag ttggctcca cagcaccacc tgagctaacc cctgtggtcc 840
aggactgcta ccatggtgat ggacagagct accgaggcac atcctccacc accaccacag 900
gaaagaagtg tcagtcttgg tcacttatga caccacaccg gcaccagaag accccagaaa 960
actacccaaa tgctggcctg acaatgaact actgcaggaa tccagatgcc gataaaggcc 1020
cctggtgttt taccacagac cccagcgtca ggtgggagta ctgcaacctg aaaaaatgct 1080
caggaacaga agcgtataaa aagctt 1106

20

<210> 14

<211> 1106

<212> DNA

<213> human

25

<400> 14

ccatgtccgt gtacctttct gaatgtaaga ctgggaatgg aaagaattac agagggacga 60

10

5
10
15

tgtccaaaac aaaaaatggc atcaccctgtc aaaaatggag ttccacttct ccccacagac 120
ctagattctc acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc 180
cagacaacga tccgcagggg ccttggtgt atactactga tccagaaaag agatatgact 240
actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg 300
gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagcccac 360
acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc 420
gtaaccccg tagggagctg cggccttggg gtttcaccac cgaccccaac aagcgtggg 480
aactttgtga catccccgc tgcacaacac ctccaccatc ttctgggtcc acctaccagt 540
gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttacctg tccgggcaca 600
cctgtcagca ctggagtga cagaccctc acacacataa caggacacca gaaaacttcc 660
cctgcaaaaa tttggatgaa aactactgcc gcaatcctga cggaaaaagg gcccctggg 720
gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct 780
ccccagtatc caggaacag ttggctccca cagcaccacc tgagctaacc cctgtggtcc 840
aggactgcta ccatggtgat ggacagagct accgaggcac atcctccacc accaccacag 900
gaaagaagtg tcagtcttgg tcattctatga caccacaccg gcaccagaag accccagaaa 960
actacccaaa tgctggcctg acaatgaact actgcaggaa tccagatgcc gataaaggcc 1020
cctggtgttt taccacagac cccagcgtca ggtgggagta ctgcaacctg aaaaaatgct 1080
caggaacaga agcgtataaa aagctt 1106

20

<210> 15

<211> 1106

<212> DNA

<213> human

25

<400> 15

ccatgtccgt ttatctgtct gactgcaaga ctgggaatgg aaagaattac agagggacga 60

	tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct ccccacagac	120
	ctagatttct acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc	180
	cagacaatga tccgcagggg ccttggtgt atactactga tccagaaaag agatatgact	240
	actgcgacat tcttgagtgt gaagaggaat gtatycattg cagtggagaa aactatgacg	300
5	gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagccac	360
	acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc	420
	gtaacccuga tagggagctg cggccttggg gtttcaccac cgaccccaac aagcgtggg	480
	aactttgtga catccccgcg tgcacaacac ctccaccatc ttctgggtcc acctaccagt	540
	gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca	600
10	cctgtcagca ctggagtga cagaccctc acacacataa caggacacca gaaaacttcc	660
	cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gccccatggt	720
	gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct	780
	ccccagtatc caggaacag ttggctccca cagcaccacc tgagctaacc cctgtggtcc	840
	aggactgcta ccatggtgat ggacagagct accgaggcac atcctctacc accaccacag	900
15	gaaagaagtg tcagtcttgg tcattctatga caccacaccg gcaccagaag accccagaaa	960
	actacccaaa tgctggcctg acaatgaact actgcaggaa tccagatgcc gataaaggcc	1020
	cctggtgttt taccacagac cccagcgtca ggtgggagta ctgcaacctg aaaaaatgct	1080
	caggaacaga agcgtataaa aagctt	1106
20	<210> 16	
	<211> 564	
	<212> DNA	
	<213> human	
25	<400> 16	
	ccatgtctga agaagaatgc augcaactga gtggagaaaa ctatgacggc aaaatttcca	60

12

agaccatgtc tggactggaa tgccaggcct gggactctca gagccacac gtcctggat 120
 acattccttc caaatttcca aacaagaacc tgaagaagaa ttactgtcgt aaccccgata 180
 gggagctgcg gccttgggtgt ttcaccaccg accccaacaa gcgctgggaa ctttgtgaca 240
 tcccccgctg cacaacacct ccaccatctt ctggccccac ctaccagtgt ctgaaggga 300
 5 cagggtgaaaa ctatgcggg aatgtggctg ttaccgtgtc cgggcacacc tgtcagcaact 360
 ggagtgcaca gaccctcac acacataaca ggacaccaga aaacttcccc tgcaaaaatt 420
 tggatgaaaa ctactgccgc aatcctgacg gaaaaagggc cccatgggtgc catacaacca 480
 acagccaagt gcggtgggag tactgtaaga taccgtcctg tgactcctcc ccagtatcca 540
 cggaacaatg gcttataaaa gctt 564

10

<210> 17

<211> 564

<212> DNA

<213> human

15

<400> 17

20

25

ccatgtctga ggaagagtgc atgcattgca gtggagaaaa ctatgacggc aaaatttcca 60
 agaccatgtc tggactggaa tgccaggcct gggactctca gagccacac gtcctgggt 120
 acattccttc caaatttcca aacaagaacc tgaagaagaa ttactgtcgt aaccccgata 180
 gggagctgcg gccttgggtgt ttcaccaccg accccaacaa gcgctgggaa ctttgtgaca 240
 tcccccgctg cacaacacct ccaccatctt ctggccccac ctaccagtgt ctgaaggga 300
 cagggtgaaaa ctatgcggg aatgtggctg ttaccgtgtc cgggcacacc tgtcagcaact 360
 ggagtgcaca gaccctcac acacataaca ggacaccaga aaacttcccc tgcaaaaatt 420
 tggatgaaaa ctactgccgc aatcctgacg gaaaaagggc cccatgggtgc catacaacca 480
 25 acagccaagt gcggtgggag tactgtaaga taccgtcctg tgactcctcc ccagtatcca 540
 cggaacaatg gcttataaaa gctt 564

<210> 18

<211> 566

<212> DNA

5

<213> human

<400> 18

10

mcatgtccga agaagaatgc atgcactgca gtggagaaaa ctatgacggc aaaatttcca 60
aggccatgtc tggactggaa tgccaggcct gggactctca gagcccacac gctcatggat 120
acattccttc caaatttcca aacaagaacc tgaagaagaa ttactgtcgt aaccccgata 180
gggagctgcg gccttggtgt ttcaccaccg accccaacaa gcgctgggaa ctttgtgaca 240
tcccccgctg cacaacacct ccaccatctt ctggtccac ctaccagtgt ctgaaggga 300
caggtgaaaa ctatcgcggg aatgtggctg ttaccgtgtc cgggcacacc tgcagcact 360
ggagtgcaca gacccctcac acacataaca ggacaccaga aaacttcccc tgcaaaaatt 420
tggatgaaaa ctactgcgcg aatcctgacg gaaaaagggc cccatggtgc catacaacca 480
acagccaagt gcggtgggag tactgtgaaga taccgtctg tgactctctc ccagtatcca 540
cggaacaatt ggcttaataa aagctt 566

20

<210> 19

<211> 854

<212> DNA

<213> human

<400> 19

25

ccatgtctga ggaagaatgc atgcactgca gtggagaaaa ctatgacggc aaaatttcca 60
agaccatgtc tggactggaa tgccaggcct gggactctca gagcccacac gctcatggat 120

14

	acattccttc caaatttcca aacaagaacc tgaagaagaa ttactgtcgt aaccccgata	180
	gggagctgcg gccttggtgt ttcaccaccg accccaacaa gcgctgggaa ctttgtgaca	240
	tcccccgctg cacaacacct ccaccatctt ctggtccac ctaccagtgt ctgaaggga	300
	caggtagaaa ctatcgcggtg aatgtggtg ttaccgtgtc cgggcacacc tgtcagcact	360
5	ggagtgcaca gaccctcac acacataaca ggacaccaga aaacttcccc tgcaaaaatt	420
	tggatgaaaa ctactgccgc aatctgacg gaaaaagggc cccatggtgc catacaacca	480
	acagccaagt gcggtgggag tactgtaaga taccgtctg tgactctcc ccagtatcca	540
	cggaacagtt ggctccaca gcaccacctg agctaacccc tgtggtccag gactgctacc	600
	atggtgatgg acagagctac cgaggcacat cctccaccac caccacagga aagaagtgtc	660
10	agtcttggtc atctatgaca ccacaccggc accagaagac cccagaaaac taccctaatg	720
	ctggcctgac aatgaactac tgcaggaatc cagatgccga taaaggcccc tgggtgttta	780
	ccacagacc cagcgtcagg tgggagtact gcaacctgaa aaaatgctca ggaacagaag	840
	cgtaataaaa gctt	854
15	<210> 20	
	<211> 854	
	<212> DNA	
	<213> human	
20	<400> 20	
	ccatgtcaga ggaagagtgc atgcactgca gtggagaaaa ctatgacggc aaaatttcca	60
	agaccatgtc tggactggaa tgccaggcct gggactctca gageccacac gctcatggat	120
	acattccttc caaatttcca aacaagaacc tgaagaagaa ttactgtcgt aaccccgata	180
	gggagctgcg gccttggtgt ttcaccaccg accccaacaa gcgctgggaa ctttgtgaca	240
25	tcccccgctg cacaacacct ccaccatctt ctggtccac ctaccagtgt ctgaaggga	300
	caggtagaaa ctatcgcggtg aatgtggtg ttaccgtgtc cgggcacacc tgtcagcact	360

5 ggagtgacac gacccctcac acacataaca ggacaccaga aaacttcccc tgcaaaaatt 420
tggatgaaaa ctactgccgc aatcctgacg gaaaaagggc cccatggtgc catacaacca 480
acagccaagt gcggtgggag tactgtaaga taccgtcctg tgactcctcc ccagtatcca 540
cggaacagtt ggctcccaca gcaccacctg agctaacccc tgtggtccag gactgctacc 600
atggtgatgg acagagctgc cgaggcacat cctccaccac caccacagga aagaagtgtc 660
agtcttggtc atctatgaca ccacaccggc accagaagac cccagaaaac tacccaaatg 720
ctggcctgac aatgaactac tgcaggaatc cagatgccga taaaggcccc tgggtgttta 780
ccacagaccc cagcgtcagg tgggagtact gcaacctgaa aaaatgctca ggaacagaag 840
cgtaataaaa gctt 854

10

<210> 21

<211> 854

<212> DNA

<213> human

15

<400> 21

ccatgtccga agaagagtgc atgcactgca gtggagaaaa ctatgacggc aaaatttcca 60
agaccatgtc tggactggaa tgccaggcct gggactctca gageccacac gctcatggat 120
acattccttc caaatttcca aacaagaacc tgaagaagaa ttactgtcgt aaccccgata 180
20 gggagctgcg gccttgggtg ttaccaccgg accccaacaa gcgctgggaa ctttgtgaca 240
tcccccgctg cacaacacct ccaccatett ctggteccac ctaccagtgt ctgaaggga 300
cagggtgaaaa ctatcgcggg aatgtggctg ttaccgtgtc cgggcacacc tgtcagcact 360
ggagtgacac gacccctcac acacataaca ggacaccaga aaacttcccc tgcaaaaatt 420
tggatgaaaa ctactgccgc aatcctgacg gaaaaagggc cccatggtgc catacaacca 480
25 acagccaagt gcggtgggag tactgtaaga taccgtcctg tgactcctcc ccagtatcca 540
cggaacagtt ggctcccaca gcaccacctg agctaacccc tgtggtccag gactgctacc 600

16

atggtgatgg acagagctac cgaggcacat cctccaccac caccacagga aagaagtgtc 660
agtcttggtc atctatgaca ccacaccggc accagaagac cccagaaaac tacccaaatg 720
ctggcctgac aatgaactac tgcaggaatc cagatgccga taaaggcccc tgggtgttta 780
ccacagaccc caqcgtcagg tgggagtact gcaacctgaa aaaatgtctc ggaacagaag 840
5 cgtaataaaa gctt 854

<210> 22

<211> 818

<212> DNA

10 <213> human

<400> 22

ccatgtcagt ttatctttca gagtgtgaaga ctgggaatgg aaagaattac agaggggacga 60
tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct cccacagac 120
15 ctagattctc acctgttaca caccctcag agggactgga ggagaactac tgcaggaatc 180
cagacaacga tccgcagggg ccctggtgct atactactga tccagaaaag agatatgact 240
actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg 300
gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagccac 360
acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc 420
20 gtaaccccgga tagggagctg cggccttggt gtttcaccac cgaccccaac aagcgtggg 480
aactttgtga catccccgcg tgcacaacac ctccaccatc ttctggtccc acctaccagt 540
gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca 600
cctgtcagca ctggagtga cagaccctc acacacataa caggacacca gaaaacttcc 660
cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gccccatggt 720
25 gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct 780
ccccagtatc cacggaacaa ttggcttaat aaaagctt 818

<210> 23

<211> 818

<212> DNA

5

<213> human

<400> 23

	ccatgtcagt gtacctttct gaatgtaaga ctgggaatgg aaagaattac agagggacga	60
	tgccaaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct cccacagac	120
10	ctagattctc acctgtctaca caccctcag agggactgga ggagaactac tgcaggaatc	180
	cagacaacga tccgcagggg ccctggtgct atactactga tccagaaaag agatatgact	240
	actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg	300
	gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagccac	360
	acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc	420
15	gtaaccccg tagggagctg cggecttggg gtttcaccac cgaccccaac aagcgctggg	480
	aactttgtga cateccccgc tgcacaacac ctccaccatc ttctggtccc acctaccagt	540
	gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca	600
	cctgtcagca ctggagtgca cagaccctc acacacataa caggacacca gaaaacttcc	660
	cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gccccatggt	720
20	gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct	780
	ccccagtatc caggaacaa ttggcttaat aaaagctt	818

<210> 24

<211> 818

25

<212> DNA

<213> human

<400> 24

	ccatgtcagt ttatctttca gagtgtgaaga ctgggaatgg aaagaattac agagggacga	60
	tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct cccacagac	120
5	ctagattctc acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc	180
	cagacaacga tccgcagggg ccctggtgct atactactga tccagaaaag agatatgact	240
	actgcgacat tcttgagtgt gaagagggaat gtatgcattg cagtggagaa aactatgacg	300
	gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagccac	360
	acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc	420
10	gtaaccocga tagggagctg cggccttggg gtttcaccac cgaccccaac aagcgtggg	480
	aactttgtga catccccgcg tgcacaacac ctccaccatc ttctggtccc acctaccagt	540
	gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca	600
	cctgtcagca ctggagtga cagaccctc acacacataa caggacacca gaaaacttcc	660
	cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gccccatggt	720
15	gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct	780
	cccagtatc caggaacaa ttggcttaat aaaagctt	818

<210> 25

<211> 818

20

<212> DNA

<213> human

<400> 25

	ccatgtctgt ttacctgtct gaatgtgaaga ctgggaatgg aaagaattac agagggacga	60
25	tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct cccacagac	120
	ctagattctc acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc	180

cagacaacga tccgcagggg ccctggtgct atactactga tccagaaaag agatatgact 240
actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg 300
gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagcccac 360
acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc 420
5 gtaaccccgga tagggagctg cggccttggt gtttcaccac cgaccccaac aagcgtggg 480
aactttgtga catccccgcg tgcacaacac ctccaccatc ttctggctcc acctaccagt 540
gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca 600
cctgtcagca ctggagtga cagaccctc acacacataa caggacacca gaaaacttcc 660
cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gccccatggt 720
10 gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct 780
ccccagtatc caggaacaa ttggttaat aaaagctt 818

<210> 26

<211> 818

15

<212> DNA

<213> human

<400> 26

ccatgtccgt ttatctttcc gagtgaaga ctgggaatgg aaagaattac agagggacga 60
20 tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct cccacagac 120
ctagattctc acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc 180
cagacaacga tccgcagggg ccctggtgct atactactga tccagaaaag agatatgact 240
actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg 300
gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagcccac 360
25 acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc 420
gtaaccccgga tagggagctg cggccttggt gtttcaccac cgaccccaac aagcgtggg 480

	aactttgtga catccccgc tgcacaacac ctccaccatc ttctgggtccc acctaccagt	540
	gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca	600
	cctgtcagca ctggagtgcg cagacccctc acacacataa caggacacca gaaaacttcc	660
	cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gcccctgggt	720
5	gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct	780
	ccccagtatc cacggaacaa ttggcttaat aaaagctt	818
	<210> 27	
	<211> 818	
10	<212> DNA	
	<213> human	
	<400> 27	
	ccatgtctgt gtacctgtcc gaatgtaaga ctgggaatgg aaagaattac agagggacga	60
15	tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct cccacagac	120
	ctagattctc acctgtctaca caccctcag agggactgga ggagaactac tgcaggaatc	180
	cagacaacga tccgcagggg ccttgggtgt atactactga tccagaaaag agatatgact	240
	actgcgacat tcttgagtgt gaagaggaat gtatgcattg cagtggagaa aactatgacg	300
	gcaaaatttc caagaccatg tctggactgg aatgccaggc ctgggactct cagagccac	360
20	acgctcatgg atacattcct tccaaatttc caaacaagaa cctgaagaag aattactgtc	420
	gtaaccccgga tagggagctg cggccttggg gtttcaccac cgaccccaac aagcgtggg	480
	aactttgtga catccccgc tgcacaacac ctccaccatc ttctgggtccc acctaccagt	540
	gtctgaaggg aacaggtgaa aactatcgcg ggaatgtggc tgttaccgtg tccgggcaca	600
	cctgtcagca ctggagtgcg cagacccctc acacacataa caggacacca gaaaacttcc	660
25	cctgcaaaaa ttggatgaa aactactgcc gcaatcctga cggaaaaagg gcccctgggt	720
	gccatacaac caacagccaa gtgcggtggg agtactgtaa gataccgtcc tgtgactcct	780

ccccagttatc caggaacaa ttggttaat aaaagctt

818

<210> 28

<211> 281

5

<212> DNA

<213> human

<400> 28

10

ccatggcagt ttatctttca gagtgtaga ctgggaatgg aaagaattac agagggacga 60
tgtccaaaac aaaaaatggc atcacctgtc aaaaatggag ttccacttct ccccacagac 120
ctagattctc acctgctaca caccctcag agggactgga ggagaactac tgcaggaatc 180
cagacaacga tccgcagggg ccctgggtgt atactactga tccagaaaag agatatgact 240
actgcgacat tcttgagtgt gaagaggaat aataaaagct t 281

15

<210> 29

<211> 345

<212> DNA

<213> human

20

<400> 29

25

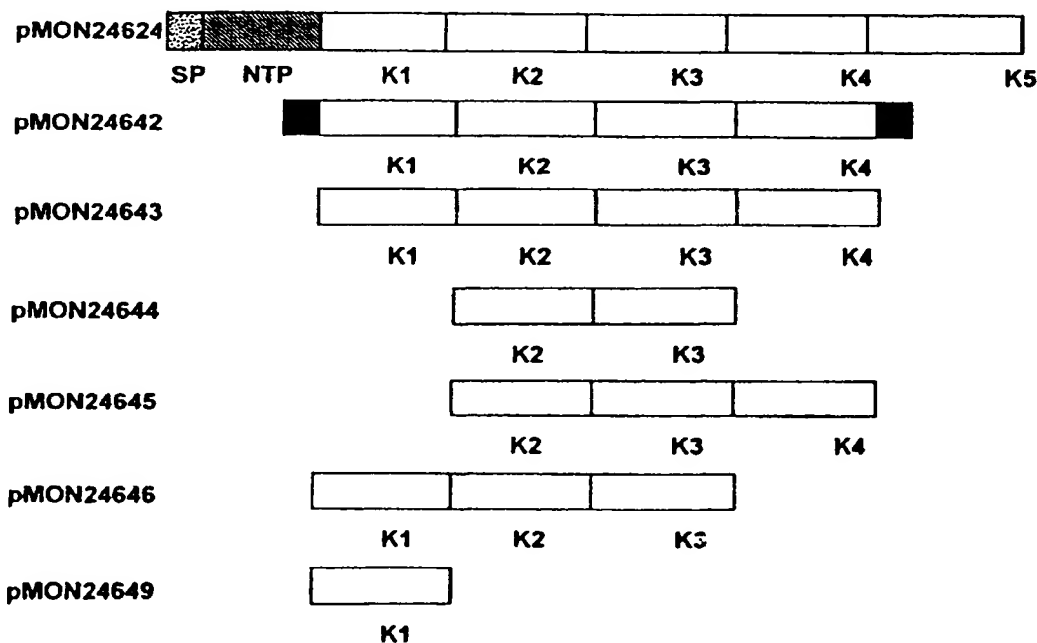
ggatccacca tgagccgcct gcccgctctg ctctgctcc aactcttgtt ccgccccgcc 60
atggccgtgt acctgtcaga atgtaagact ggaatggaa agaattacag agggacgatg 120
tccaaaacaa aaaatggat cacctgtcaa aaatgggtt ccacttctcc ccacagacct 180
agattctcac ctgctacaca cccctcagag ggactggagg agaactactg caggaatcca 240
gacaacgac cgcagggggc ctggtgctat actactgatc cagaaaagag atatgactac 300
tgcgacattc ttgagtgtga agaggaataa taaaagcttg gattc 345



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 9/68, C07K 1/113		A3	(11) International Publication Number: WO 99/35248
			(43) International Publication Date: 15 July 1999 (15.07.99)
(21) International Application Number: PCT/US99/00048 (22) International Filing Date: 11 January 1999 (11.01.99) (30) Priority Data: 60/071,247 12 January 1998 (12.01.98) US (71) Applicant (for all designated States except US): G.D. SEARLE & CO. [US/US]; Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): VIOLAND, Bernard, N. [US/US]; 1710 Woods Bend Lane, Glencoe, MO 63038 (US). POLAZZI, Joseph, O. [US/US]; Chesterfield, MO (US). CASPERSON, Gerald, F. [US/US]; Ballwin, MO (US). (74) Agents: WILLIAMS, Roger, A. et al.; G.D. Searle & Co., Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> (88) Date of publication of the international search report: 11 November 1999 (11.11.99)	

(54) Title: METHOD OF REFOLDING ANGIOSTATIN



(57) Abstract

An efficient method for expressing, solubilizing, refolding, and purifying, angiotensin from inclusion bodies expressed in bacteria is disclosed. Nucleic acids encoding full-length angiotensin and subfragments encoding kringle domains (K1-4, K2-3, K2-4, K1-3, and K1) are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No

PC 1/US 99/00048

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N9/68 C07K1/113

According to international Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	YIHAI CAO ET AL.: "Kringle domains of human angiostatin" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 271, no. 46, 15 November 1996 (1996-11-15), pages 29461-29467, XP002086445 MD US cited in the application	1-3, 19-22, 33,34, 37-40, 45-47
Y	page 29462, right-hand column, paragraph 3 --- -/-	4-18, 23-32, 35,36, 41-44,48

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

28 April 1999

Date of making of the international search report

24.09.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx: 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

MONTERO LOPEZ B.

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 99/00048

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SCOTT CLEARY ET AL.: "Purification and characterization of tissue plasminogen activator kringle-2 domain expressed in Escherichia coli" BIOCHEMISTRY, vol. 28, no. 4, 21 February 1989 (1989-02-21), pages 1884-1891, XP002101498 EASTON, PA US	1-3,15, 16, 19-22, 33,34, 45-47
Y	page 1885, left-hand column, paragraph 5 -right-hand column, paragraph 1; figure 2 page 1886, right-hand column, paragraph 2 -page 1887, right-hand column, paragraph 1 ---	4-14,17, 18, 23-32, 35,36, 41-44,48
X	EP 0 523 296 A (ELI LILLY AND COMPANY) 20 January 1993 (1993-01-20)	1-9, 11-14, 19,20, 29-32
Y	page 3, line 18 - line 21; examples 34,36 ---	10, 15-18, 23-28, 35,36, 41-44,48
Y	BERNHARD FISCHER ET AL.: "Isolation, renaturation, and formation of disulfide bonds of eukaryotic proteins expressed in Escherichia coli as inclusion bodies" BIOTECHNOLOGY AND BIOENGINEERING INCLUDING: SYMPOSIUM BIOTECHNOLOGY IN ENERGY PRODUCTION AND CONSERVATION., vol. 41, no. 1, 5 January 1993 (1993-01-05), pages 3-13, XP000605146 NEW YORK US cited in the application page 4, left-hand column, paragraph 2 - last paragraph page 5, left-hand column, paragraph 2 -page 9, right-hand column, paragraph 1; table VI ---	4-9, 11-18, 23-32, 35,36, 41-44
Y	WO 97 41824 A (ABBOTT LABORATORIES) 13 November 1997 (1997-11-13) abstract; figure 3 -----	10,48

INTERNATIONAL SEARCH REPORT

International application No

PCT/US 99/ 00048

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons

1. ☐ Claims Nos. :
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos. :
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically
3. ☐ Claims Nos. :
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows.

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos. :
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos. :

1-48

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest
- ☐ No protest accompanied the payment of additional search fees

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99/00048

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-48

Method for producing angiostatin including a step of refolding at an elevated pH

2. Claim : 49

Polypeptide encoded by a sequence selected from SEQ ID NO:10 through 29.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/00048

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 523296	A	20-01-1993	CA 2047063 A	16-01-1993
			JP 5076360 A	30-03-1993

WO 9741824	A	13-11-1997	US 5801146 A	01-09-1998
			AU 3060697 A	26-11-1997
			CZ 9803426 A	12-05-1999
			EP 0910571 A	28-04-1999
